NRC

Sound Localization by Human Observers Symposium Proceedings

AD-A212 877

89 9 26 171

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS N. A.
2a SECURITY CLASSIFICATION AUTHORITY N. A. 2b DECLASSIFICATION / DOWNGRADING SCHEDULE N. A.	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S) XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION NAS-National Research Council 6b OFFICE SYMBOL (If applicable) CHABA	7a NAME OF MONITORING ORGANIZATION Office of Naval Research
6c ADDRESS (City, State, and ZIP Code) 2101 Constitution Ave., N. W. Washington, DC 20418	7b ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000
8a NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b OFFICE SYMBOL (If applicable) Code 1142PS	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N0001487C0342
8c ADDRESS (City, State, and ZIP Code) Arlington, VA 22217-5000	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 61153N 42 PROJECT NO: RR04209 TASK NO: RR0420901 WORK UNIT ACCESSION NO: 4426124

11 TITLE (Include Security Classification)

(U) Sound localization by human observes: Symposium Proceedings

12 PERSONAL AUTHOR(S)
CHABA

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM 88/10/01 TO 89/09/30	89/07/01	57

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Sound localization, binaural perception in space and in headphones, head-related trasnfer function, sound sources.

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

(U) The relation between binaural perception of actual sounds in space and of stimuli presented over headphones, based on head-related transfer functions, showed a high correlation. Listeners reported that the sound image presented over headphones remained internal to the head and not external in space, as with natural sound. Head movement, context, bias, experience, and integration of information across senses remain probable contributions to externalizing headphone-delivered sounds at their real-world positions. Sensory information provided by one sense modality may modulate that received by another modality via motor and/or efferent control. Interest was expressed in the possibility of constructing, and adapting to, super localization systems. Symposium focused on mammalian research but crucial insights were reported on other animal models. Thus there was demonstration that auditory space for the cat may be coded by spatial maps within the superior colliclus. Some human data were presented on the poorer accuracy of locating a source in the presence of multiple sources. Monaural localization, & the coding of sound movement were discussed.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL John J. O'Hare	22b TELEPHONE (Include Area Code) (202) 696-4502 22c OFFICE SYMBOL Code 1142PS

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Sound Localization by Human Observers Symposium Proceedings

Committee on Hearing, Bioacoustics, and Biomechanics
Commission on Behavioral and Social Sciences and Education
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1989

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Frank Press is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Robert M. White is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Samuel O. Thier is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Frank Press and Dr. Robert M. White are chairman and vice-chairman, respectively, of the National Research Council.

This work relates to U.S. Department of the Navy Contract N00014-87-C-0272 issued by the Office of Naval Research under Contract Authority NR 201-124. However, the content does not necessarily reflect the position or the policy of the Department of the Navy or the government, and no official endorsement should be inferred.

Available from:
Committee on Hearing, Bioacoustics, and Biomechanics
National Research Council
2101 Constitution Avenue, N.W.
Washington, DC 20418

Printed in the United States of America

COMMITTEE ON HEARING, BIOACOUSTICS, AND BIOMECHANICS

JOHN H. MILLS (Chair), Department of Otolaryngology, Medical University of South Carolina
ROBYN M. COX, Speech and Hearing Center, Memphis State University
ROBERT A. DOBIE, Department of Otolaryngology, University of Washington
C. DANIEL GEISLER, Department of Neurophysiology, University of Wisconsin
IRA J. HIRSH, Washington University and Central Institute for the Deaf, St. Louis
WILLIAM M. RABINOWITZ, Research Laboratory of Electronics, Massachusetts Institute of Technology
JAMES C. SAUNDERS, Department of Otorhinolaryngology and Human Communications, University of Pennsylvania
ROBERT D. SORKIN, Department of Psychology, University of Florida
CONSTANTINE E. TRAHIOTIS, Department of Otolaryngology, University of Connecticut Health Center
DIANNE J. VAN TASELL, Department of Communication Disorders, University of Minnesota
HENNING E. VON GIERKE, Yellow Springs, Ohio

MILTON A. WHITCOMB, Study Director
ARLYSS K. WIGGINS, Administrative Secretary

CONTRIBUTORS

MARIANNE ARRUDA, University of Wisconsin
ROBERT BERKOVITZ, Sensimetrics Corp., Cambridge, Mass.
JENS BLAUERT, Ruhr University, D-Bochum, Federal Republic of Germany
ROBERT BUTLER, University of Chicago
H. STEVEN COLBURN, Boston University
NATHANIEL I. DURLACH, Massachusetts Institute of Technology
RAYMOND H. DYE, Loyola University
MARK A. ERICSON, Wright-Patterson Air Force Base, Ohio
SCOTT H. FOSTER, Crystal River Engineering, Groveland, Calif.
D. WESLEY GRANTHAM, Wilkerson Speech and Hearing Center, Nashville, Tenn.
DAVID M. GREEN, University of Florida
ERVIN HAFTER, University of California, Berkeley
WILLIAM M. HARTMANN, Michigan State University
RICHARD HELD, Massachusetts Institute of Technology
IRA J. HIRSH, Washington University and Central Institute for the Deaf, St. Louis
DORIS KISTLER, University of Wisconsin
GEORGE F. KUHN, Vibrasound Research Corporation, Englewood, Colo.
SHIGEYUKI KUWADA, University of Connecticut Health Center
JAMES R. LACKNER, Brandeis University
RICHARD L. MCKINLEY, Wright-Patterson Air Force Base, Ohio
JOHN C. MIDDLEBROOKS, University of Florida
DAVID R. PERROTT, California State University, Los Angeles
CHRISTOPH POESSELT, Ruhr University, D-Bochum, Federal Republic of Germany
EDGAR A.G. SHAW, National Research Council of Canada, Ottawa, Ont.
ROBERT D. SORKIN, University of Florida
RICHARD M. STERN, Carnegie Mellon University
CONSTANTINE TRAHIOTIS, University of Connecticut Health Center
CHARLES S. WATSON, University of Indiana
ROBERT B. WELCH, University of Kansas
ELIZABETH M. WENZEL, Ames Research Center, National Aeronautics and Space Administration
FREDERIC L. WIGHTMAN, University of Wisconsin
WILLIAM A. YOST, Loyola University
GREG L. ZACHARIAS, Charles River Analytics, Inc., Cambridge, Mass.
PATRICK M. ZUREK, Massachusetts Institute of Technology

Foreword

The Committee on Hearing, Bioacoustics, and Biomechanics (CHABA) of the National Research Council was asked by the Air Force Office of Scientific Research to plan and conduct a symposium on the current research concerning sound localization by human observers. The symposium, held at the National Academy of Sciences October 14-16, 1988, began with 10 papers which provided reviews and overviews of some of the basic data and theories concerning auditory localization. Two individuals spoke about each topic: one who presented a review and overview of the topic and a second who provided comments about the first speaker's presentation. A number of observers were invited to the symposium, and they were asked to stimulate discussion on each topic. The symposium ended with 11 talks on current research that either were aimed at direct applications of current knowledge about localization or were closely linked to application. In addition to the presentations, a number of the speakers provided tape recordings of some of their work.

Although many talks were presented over the three days of the symposium, a number of topics on auditory localization were not included. We were unable to include information on such topics as research on masking of binaurally presented stimuli, developmental aspects of localization, and animal models of localization. The topics chosen deal directly with human localization of sounds, especially if the topic might be relevant to the application of current knowledge. Some of the topics provide a background on human psychophysics and localization, in order to present a context for some of the topics on current research. We also wanted to concentrate on laboratory research that has appeared in the scientific literature rather than on direct applications involving products, devices, or marketable procedures. Although there are a variety of products now on the market and certainly more to come, our aim was to provide a scientific review of localization as it might pertain to current and future applications.

This report consists of a brief summary of the proceedings and abstracts of the talks presented at the symposium. On behalf of CHABA, I would like to thank the members of the organizing committee—Nathaniel Durlach, Ira J. Hirsh, Charles S. Watson, and Frederic Wightman—not only for organizing the symposium but also for their efforts in drafting the summary.

William A. Yost, *Chair*
Symposium on Sound
Localization by
Human Observers

Contents

Part II
Applications of Localization Data and Theories

OTHER APPLICATIONS

Summary

Colin Cherry often expressed wonderment that although we have two ears we hear only one world. This fascination with human sound localization existed prior to Cherry's observation and has continued ever since. Despite a rich history of research, technical limitations have slowed the ability to answer many compelling questions. The Committee on Hearing, Bioacoustics, and Biomechanics (CHABA) Symposium on Sound Localization by Human Observers revealed that many of these limitations are being lifted and that the field appears to be at the threshold of an explosion of new data, theories, and applications. Chief among the new advances are the miniaturization of wideband, high-fidelity microphones and the ability of small computers to provide digital acoustical processing of complex signals in a short period of time. Accompanying the technological advances are a number of studies that show the potential of these technologies for enriching our knowledge of sound localization. Furthermore, the advances made in technology and in the laboratory have paved the way for the application of this knowledge to a wide range of practical problems, including those concerned with the use of auditory space in virtual environments, music recordings, and aids for the hearing impaired.

A major question that permeated the symposium concerned the relation between binaural perception of actual sounds in space (localization) and binaural perception of stimuli presented over headphones (lateralization). Newer techniques, which use the head-related transfer function (HRTF) measured from a subject's or manikin's ear canals to filter stimuli before presentation over headphones, demonstrate both the usefulness of headphone studies and the potential of such studies to answer basic questions and provide important applications. There appears to be a high correlation between the judgments made by listeners presented actual sound sources and those made for headphone-delivered stimuli based on HRTF simulations (especially in the horizontal plane). Because the HRTFs contain most, if not all, of the purely acoustical information at a listener's ears relevant to sound source position, experimenters can use alterations in the HRTFs and headphone-presented stimuli to study a wide range of problems that have been difficult to investigate when sounds had to be presented over loudspeakers. It appears, therefore, that data collected in experiments that use headphones will continue to provide valid information for understanding human localization.

Prior to using the HRTF to filter stimuli, almost all listeners reported that the sound image presented over headphones remained internal to the head and not external in space, as with a natural sound. Head movements had been assumed to be a major variable in

externalizing sounds. Now, the experiences of listeners presented HRTF simulations over headphones indicate that many externalize the sound. These reports suggest that head movement may not be as necessary as previously theorized. However, a careful study of the externalization of sound images with headphone-delivered HRTF simulations has not been completed. In addition, anecdotal references suggest that not all listeners agree that the headphone-delivered HRTF simulations appear to be externalized: some subjects require experience with the presentations before they form an externalized image, while other listeners who externalize the image place all images behind them even though the simulations presented the sounds from in front. Thus, variables such as head movement, context, bias, experience, and integration of information across senses remain probable contributors to externalizing headphone-delivered sounds at their real-world positions. With advances such as the use of the HRTF, additional work on the question of externalization can now be undertaken. Such research may reveal some crucial insights into this classic problem of auditory perception.

It has long been recognized that locating a sound in space involves a number of sense modalities, including auditory, vestibular, visual, and kinesthetic. The integration of these sensory inputs is clearly crucial for sound localization applications. Sensory integration and localization behavior can also be influenced by the experience of the listener in both normal and altered environments. The sensory systems often adapt to one environment, producing behaviors that may differ from those measured in other environments. In addition, sensory information provided by one sense modality may modulate that received by another modality via motor and/or efferent control. Thus, sound localization behavior may be modifiable and is not limited to the auditory system. A firmer understanding of sensory integration and adaptation is required to understand better the entire process of sound localization and to explore fully practical applications. Of particular interest is the possibility of constructing, and adapting to, super localization systems. These are systems in which localization cues are magnified and localization performance is superior to normal performance.

Accompanying the advances in measuring sounds at the ears and in understanding localization perception have been equally important new physiological findings. The symposium focused on mammalian research, but the presenters recognized the crucial insights gained from other animal models. The demonstration that auditory space for the cat may be coded by spatial maps within the superior colliculus (SC) raises the question of whether auditory spatial maps can be discovered in other neural centers. Visual space and somatosensory space are coded via spatial maps within the SC, and the SC plays a dominant role in coordinating head and body movement with sensory input to aid in tracking sources in the environment. If information about auditory space is to be integrated with that from other senses within the SC, then it is logical that auditory space be coded in the same way as other spatial information, that is, by spatial maps. However, since auditory spatial maps have not been found outside the superior colliculus in either cats or a few other species, the question of the form of the central code for auditory space is still to be answered. The difficulty of studying the central nervous system physiology was discussed at the symposium, especially in regard to the role anesthesia may play in increasing the inhibition of neural activity. Both the question of the exact way in which interaural time and level differences are processed in the brain stem and the way in which this processed neural information is sorted in the higher brain centers (i.e., via spatial maps or some other form) are being studied in many laboratories. The ability now to make accurate sound measurements at animals' ears is as much an aid to the physiological experimenter as it is to the physical and psychophysical acoustician.

The vast majority of the data and theories concerning localization relate to single sound

sources or, at most, to two sound sources. However, in most natural situations a number of sound sources are present and the listener must process the information from many, if not most, of these multiple sources. A great deal more needs to be known about localizing multiple sound sources if the current applications are to be maximally successful. For instance, some current data suggest that humans are not as accurate at locating one source in the presence of many sources as they are when they localize the target source in isolation. A few suggestions were presented at the symposium that describe ways of processing signals such that target location in a multisource environment might be as accurate as target location of a single source. However, these suggested techniques are difficult to evaluate in the absence of data.

A number of other questions were posed at the symposium. One concerned the extent to which listeners can locate sounds monaurally. Another concerned whether or not movement of a sound in space is coded directly in the auditory system, as visual motion is in the visual system, or whether auditory motion is derived from cues of time and distance. There are data suggesting that the auditory system is slow to process binaurally changing stimuli, at least relative to the time it takes to process other types of sound. A clear relationship between the slow processing time for changing binaural variables and the more direct measures of auditory movement has not been established. A further issue, discussed in a number of presentations, concerned the localization of sounds in environments producing reflections. Phenomena associated with the onsets of direct and reflected sounds, as well as between the steady-state portion of the sounds, are often confused when topics such as the precedence effect and the Haas effect are discussed. The effects of the various types of reflections on many aspects of auditory perception (e.g., localization, sound quality, recognition) have not yet been adequately clarified.

Models of auditory localization and differences from listener to listener are two topics that were common to many presentations. Most current models of localization are, in fact, based on lateralization data, and suggest that some form of cross-correlation of the output of tuned channels at each ear provides a good model for a great deal of these data. It was recognized that these models need to be tested on data from localization studies and the modelers need to consider the data acquired from the physiological exploration of localization. These considerations may suggest new models as well as variations in existing models. Large individual differences are obtained in the HRTFs measured from different people as well as from the two ears of the same person; however, the functional significance of such differences is not yet clear. Descriptions of the actual causes of the physical difference among the HRTFs are also still to be determined. In addition, some of the presentations indicated that significant individual differences exist in a number of localization tasks, especially when such tasks are performed in altered environments and/or involve listeners adapting to one environment or another.

This summary only highlights a number of the issues presented and discussed over the three-day symposium. Many more questions than answers were provided, but the new tools and the creative ways of using them show promise that many of these questions will soon be answered and new ones proposed. It was also apparent that applications will be occurring in a number of areas, even though there is still much that is unknown. Some of the applications are already successful in providing localization information for single sound sources in restricted environments. It is clear that new applications will quickly move beyond these early successes. As these applications are tested, the failures and the successes may provide additional information about auditory localization.

Part I
Reviews of Basic Research and Comments

Spatial Hearing

DAVID M. GREEN

For frequencies below about 20,000 Hz, the propagation of acoustic energy in most ear canals can be regarded as a plane progressive wave. Thus, a complete description of the proximal stimulus for single-eared listening can be described as a function of one variable, for example, the variation of sound pressure as a function of time, p(t). A binaural stimulus can, therefore, be described as two pressure waveforms, $p_1(t)$ and $p_2(t)$. Differences in these two waveforms allow the listener to make inferences about the acoustic features of the environment surrounding the listener. Among these features are the location of the source, its distance, and the sound field itself—the size and structure of the sound enclosure. Among the more critical aspects of the sound field are the reflections and refractions of those objects near the ear canals, such as the listener's body, head, and external ears.

It is a small miracle that any valid inference can be made about the nature of the source from these data, since so many features of the source and the listener's acoustic environment influence these two pressure waves. One of the major problems is that most everyday surfaces produce stong acoustic reflections. These hard surfaces produce virtual sources that mimic the original acoustic source in practically all aspects, except location. Yet, in most circumstances, we hear with considerable clarity a single acoustic source, located at a definite position in space. How we accomplish this complicated pattern of inferences, what we might call with Jens Blauert, *Spatial Hearing*, is the topic of this volume.

Spatial hearing is a topic very different from what we generally study using conventional earphone listening. It is certainly a more complicated process because of the multiple cues present in the stimulus and the need to resolve a considerable amount of ambiguous information. Some years ago, Lord Rayleigh suggested that the listener might rely primarily on two stimulus differences: interaural time and intensity. With the usual analytic reductions common to most science, the study of these two cues with earphones has occupied the bulk of our efforts in binaural hearing for the past half-century. Often, understanding the relative contributions of these two cues appears, regrettably, to be the sole aim of some earphone reasearch. For most topics in spatial hearing, one can make a strong argument that earphone experiments should only be used to check some hypothesis developed to account for some phenomenon observed in free-field listening.

It is my sincere hope that this volume will initiate a more concentrated effort to study the fascinating mechanisms that underlie hearing without earphones. It is my conviction that, paradoxically, the understanding of these processes will ultimately provide the information

we need to produce earphones of genuine worth, and that among the ultimate benefits will be more useful electronic amplification (hearing aids) for listeners with impaired hearing.

Comment: Psychoacoustic Studies of Auditory Space

IRA J. HIRSH

It is certainly the case, as Green points out, that a large portion of research on auditory localization has concerned lateralization and the interaural cues associated with judgments of azimuth on the horizontal. Auditory space has not been neglected, however; and even the earliest studies were devoted to the exploration of real, external space.

From Sylvanus Thompson to Lord Rayleigh, to Stevens and Newman, to Mills, investigators have sought to know about the accuracy of listeners in localizing sources in external space—most particularly in the horizontal plane. Then, from the early 1900s, rubber tubes and later earphones were employed to test hypotheses about the cues that were most likely giving rise to these abilities in real space. Eventually, we could know that a few degrees of displacement enabled discrimination, but that result depended on the standard location (most favorable is directly in front) and also on frequency. Correlated with that result was the minimum discriminable interaural time disparity of 20 microseconds that emerged from von Hornbostel and Wertheimer, in addition to the value of 600 microseconds that represented a sound image at the side (90 or 270 degrees). Intensity differences were somewhat less salient, but different estimates of a trading relation were subsequently given.

Distance or depth and elevation have not been attended to, partly since they are not easily handled by interaural cues. Bekesy emphasized the ratio of direct to reflected sound for distance. The role of head movement for vertical localization was elegantly evaluated by Wallach.

In the past 50 years there has emerged an interest in two related aspects of real space—the singularity of apparent locations of sound sources and reliable localizations of those sources in spite of confounding reflections. Even there, however, earphone studies like those of Wallach, Newman, and Rosenzweig brought hypothesized cues under careful control, with the result that previous explanations of auditory perspective (e.g., Steinberg and Snow) had to be changed.

We do indeed have more information about lateralization than about auditory space. Echo suppression, correlations between the waveforms of signals at the two ears, auditory-vestibular-motor aspects of the role of head movements, and other phenomena provide a fertile background for future studies.

REFERENCES

Green, D.M.
1988 Audition: Psychophysics and perception. Pp. 366-371 in Atkinson, Hernstein, Lindzey, and Luce, eds., *Stevens' Handbook of Experimental Psychology*, Vol. 1. New York: Wiley.

Mills, A.W.
1972 Auditory localization. Chapter 8 in J.V. Tobias, ed., *Foundations of Modern Auditory Theory*. New York: Academic Press.

Wallach, H.
1940 The role of movements and vestibular and visual cues in sound localization. *Journal of Experimenatal Psychology* 27:339-368.

Wallach, H., E.B. Newman, and M.R. Rosenzweig
1949 The precedence effect in sound locatization. *American Journal of Psychology* 62:215-336.

Physical Measurements and Models Pertaining to the Torso, Head, and External Ear

GEORGE F. KUHN

As sound is transmitted from the field to the eardrum, information about the source location is encoded onto the acoustic signal. This encoded information results from the path-length differences between the ears, from the diffraction of the sound by the head and torso, from resonances of the external ear, and from relative wave motion within the pinna and the ear canal. The amplitude and phase distortions, or time delays in the case of pulsed transmission, encoded onto the signal by the physical components, such as the head, the torso, the pinna, and the ear canal, depend on the frequency and on the angle of incidence. Measurement methods and analytical models for measuring or predicting these distortions, which form the localization cues, are described. Analytical and experimental models, which have been used in the past to provide a better understanding of the mechanisms that generate specific localization cues in different frequency regions are presented.

A sorting of the effects that the individual physical components have on the acoustic signal shows that different physical components become acoustically prominent at (approximate) successive octave intervals. Below 1 kHz, the amplitude and phase distortions result from the frequency-dispersive wave motion around the head and from backscattering of the sound by the torso. Thus, below 1 kHz the head and torso generate the localization cues. Above approximately 1 kHz, the fundamental resonance of the external ear begins to shape the pressure response at the eardrum. Above approximately 2 kHz, the pinna becomes directional, necessitating an accounting for the pinna's gross anatomy. Above approximately 4 kHz, the major anatomical features of the ear canal and the closeness of the acoustical coupling between the ear canal and the pinna shape the pressure response at the eardrum. Also, relative wave motion within the pinna begins to take place, producing directional peaks and nulls in the pressure response at the eardrum. Thus, the major anatomical features of the pinna (concha, fossa, and helix) and of the ear canal need to be accounted for. Above approximately 8 kHz, relative wave motion within the smaller pinna features and within the ear canal plays an important role in the transformation of the field pressure to the eardrum. Therefore, a detailed model of the anatomical features of the pinna and the ear canal is required at very high frequencies. As a result, at each successive octave, additional physical components or features must be included in the measurements or mathematical analyses. Measurements and analytical results that describe these amplitude and phase distortions are presented for simple geometrical configurations in these different frequency regimes. These reduced models provide an overview of the typical acoustical

effects of the different physical components on the field-to-eardrum transformations and on the localization cues

The effects of geometrical perturbations in these physical components illustrate that there is also a fine structure in these transformations that exceeds the minimum detectable values. How much of this fine structure in the signal, which varies between individuals, between ears, and with time as the anatomy changes, provides personalized localization cues remains to be determined.

Comment: Directionality and Modal Characteristics of the External Ear

EDGAR A.G. SHAW

Measurements of the sound pressure level (SPL) made at the center of an earplug sealing the ear canal entrance (the blocked meatus condition) and measurements made at the eardrum show that there is almost an identical directionality up to approximately 15 kHz. When such measurements are performed with a special progressive wave source at grazing incidence close to the ear, it is possible to obtain precise information about the high-frequency characteristics of the external ear operating essentially in isolation from the torso. Measurements on 10 real human ears and 9 replicas indicate that all adult human ears share certain major acoustical characteristics; there is especially a substantial increase in response (~10 decibels [dB]) between 5 and 10 kHz when the source elevation is increased from 0 to 60 degrees. This and other high-frequency characteristics are determined by the normal modes of the external ear. Geometrical models in which the frequencies, pressure distributions, directionalities, and excitation factors of these modes are well matched to those of the average ear have response curves that closely resemble those of real human ears. In practice, this means matching six modes whose frequencies are approximately 4.3, 7.1, 9.6, 12.1, 14.4, and 16.7 kHz. (The presence of the open ear canal increases the number of modes from 6 to 8.)

The differences between the response curves of human ears are as impressive as the similarities. In some ears, for example, there are deep minima in the response curves and very large variations in response with frequency and angle of incidence (~30 dB). In others, the minima are few in number and the dependence of response on source elevation is much less pronounced. It seems likely that these characteristics are closely associated with the accuracy of sound localization in the vertical direction.

Finally, measurements on supposedly matched pairs of pinnas show that even apparently minor differences in pinna geometry, such as the acoustical connection between fossa and cymba and the openness of the concha, strongly affect the model characteristics and directionality patterns of the ear.

REFERENCE

E.A.G. Shaw

1982 External ear response and sound localization. In R.W. Gatehouse, ed., *Localization of Sound.* Groton, Conn.: Amphora Press.

Localization of Sound in Space

ROBERT BUTLER

This review of localization of sound in space is restricted to those studies on normal-hearing subjects. First, the historical background of those general theories that have survived the past several decades is briefly discussed, culminating in the classic 1936 paper of Stevens and Newman. Despite the fact that the field was preempted at this point by research on lateralization of sound, enough new localization data are available to supplement the duplex theory of localization. Much of these data center around the role of spectral cues in localization—cues that are furnished in large part by the pinna. Results of experiments are presented on localization of sound in the median sagittal plane. Here, binaural difference cues are minimal. Spectral cues are also critical for monaural localization; distort the pinna and performance deteriorates. A number of studies relating to this topic are covered. The problem of front/back discrimination of sound sources is addressed in the context of spectral cues. The influence of stimulus bandwidth on monaural localization proficiency is documented, and those frequency segments of the sound's spectrum that bias monaural location judgments are also shown to bias binaural judgments similarly, notwithstanding the presence of binaural difference cues in the latter situation. A possible relation between the apparent location of a sound and the features of its spectrum is discussed next. Included in this discussion are Blauert's directional bands and our maps of spatial referents. Lastly, those data on the precedence effect generated by free-field studies are incorporated into the overall treatment of sound localization.

Comment: Localization of Sound in Space

NATHANIEL I. DURLACH

Most studies concerned with the ability to identify sound-source direction ignore the tendency for resolution between two fixed directions to decrease as the total range of direction included in the stimulus set increases. This failure to take proper account of stimulus range leads to inappropriate comparisons of different data sets and to inappropriate theoretical models. These comments remind investigators of this effect (discussed previously in the localization literature by Serle and colleagues) and present some new data on this effect for the case of interaural time delay.

Phenomenology of Sound Image Lateralization

H. STEVEN COLBURN

This review covers the phenomenology of sound image lateralization. The lateralization of sound image is the formation of a judgment about the internal location of an auditory image. This is usually interpreted to be a report of the location of a discrete image along the interaural axis. This lateral position judgment is a fundamentally subjective phenomenon that is difficult to quantify with confidence and consistency. Although lateralization is obviously related to the localization of physical sound sources, a process that we all experience in our interaction with the world of acoustic sources, lateralization is based on internal images that are generated by the unnatural stimulus condition in which each ear is stimulated by a separately generated waveform.

The auditory images that give rise to lateralization judgments may be extremely complex, and the lateral position of an image is often a poorly defined variable. This ambiguity is perhaps a consequence of the complexity and implausibility of the physical sources that could give rise to such stimuli. In any case, one should consider the nature of the auditory image in studies of lateralization. The perceptions may be dominated by the extent of the image, by the multiplicity of the subimages, by the image shape, or by other complexities.

This review proceeds from simple to complex stimuli, generally from stimuli with few to those with many degrees of freedom. For each stimulus, we discuss the nature of the image, the dependence of the nature and location of the image on the physical parameters of the stimulus, and the influence of the presence of other stimuli, either simultaneous or sequential. Some consideration is given to quantitative descriptions of these phenomena, but no generative or mechanistic models are presented or discussed.

The physical parameters of the stimuli must be carefully defined to avoid confusion and ambiguity. For example, the simple concept of interaural time delay must be carefully delineated to allow discussion of onset and offset effects for burst stimuli and to distinguish ongoing envelope differences from ongoing fine-structure differences. Similarly, the interaural intensity difference must be defined over a specific interval of time and, as such, is a function of time over the duration of the stimulus (unless the defining interval is taken to be the duration of the stimulus).

Our review first considers stimuli that are restricted in time and/or frequency, such as clicks or tone bursts, and proceeds to stimuli with greater degrees of freedom, such as modulated tones, repeated clicks, and narrowband noises. Even in the simplest cases, there are situations that have ambiguous or multiple lateral positions and for which perceptual

attributes such as image width may dominate the perceptions. Examples include phase differences near the antiphasic point, unnatural combinations of interaural time and intensity differences, and other contradictory combinations of parameters.

For stimuli with many degrees of freedom, perceptions are often better characterized by multiple perceptual objects. For example, stimuli that cover disjointed regions of time and/or frequency are often perceived as distinct objects. These objects, although able to be identified with different parts of the physical waveforms, often show perceptual interactions, and thus, lateralization judgments with the whole waveform are often different than judgments made with the separate parts of the waveform. The mutual interactions of the separate images include pulling effects, repulsion effects, and masking effects.

Various descriptions of sound image lateralization are listed below and are described in Figures 1 to 5.

LATERALIZATION OF PURE TONES (NO ONSETS OR OFFSETS)

1. Interaural phase effects
   - Only sensitive below about 1,500 Hz
   - Cyclic position curves
   - Maximum image displacement near 90 degrees
   - Ambiguity near 180 degrees
   - Constant laterality for constant time?
   - Binaural beats
2. Interaural amplitude effects
   - Sensitive at all frequencies
   - Fully lateralized by 12 dB
   - But sensitive to increments at 40 dB
   - Little frequency dependence
3. Combination of phase (time) and amplitude
   - Time and intensity trading
   - Lateral position versus phase curves are displaced upward with intensity difference
4. Images are perceptually complex
   - Multiple images are perceived
   - Image shapes vary with stimulus parameters
   - JND performance better than that predicted for lateral position alone

ONSET AND OFFSET DELAY EFFECTS

1. Both onset and offset affect lateralization
   - Can trade onset and ongoing differences
   - Onset effects are stronger for short durations and short rise times
   - Some effects even with 200 ms rise times
2. Onset-offset effects at all frequencies
3. Transient stimuli (clicks) are onsets
   - Time-intensity trading applies to clicks
   - Filtered clicks show trading at all frequencies
4. For long-duration tone bursts with short rise times, one hears a click at the start of the burst
5. Onset time JNDs are very sensitive to level and rise time

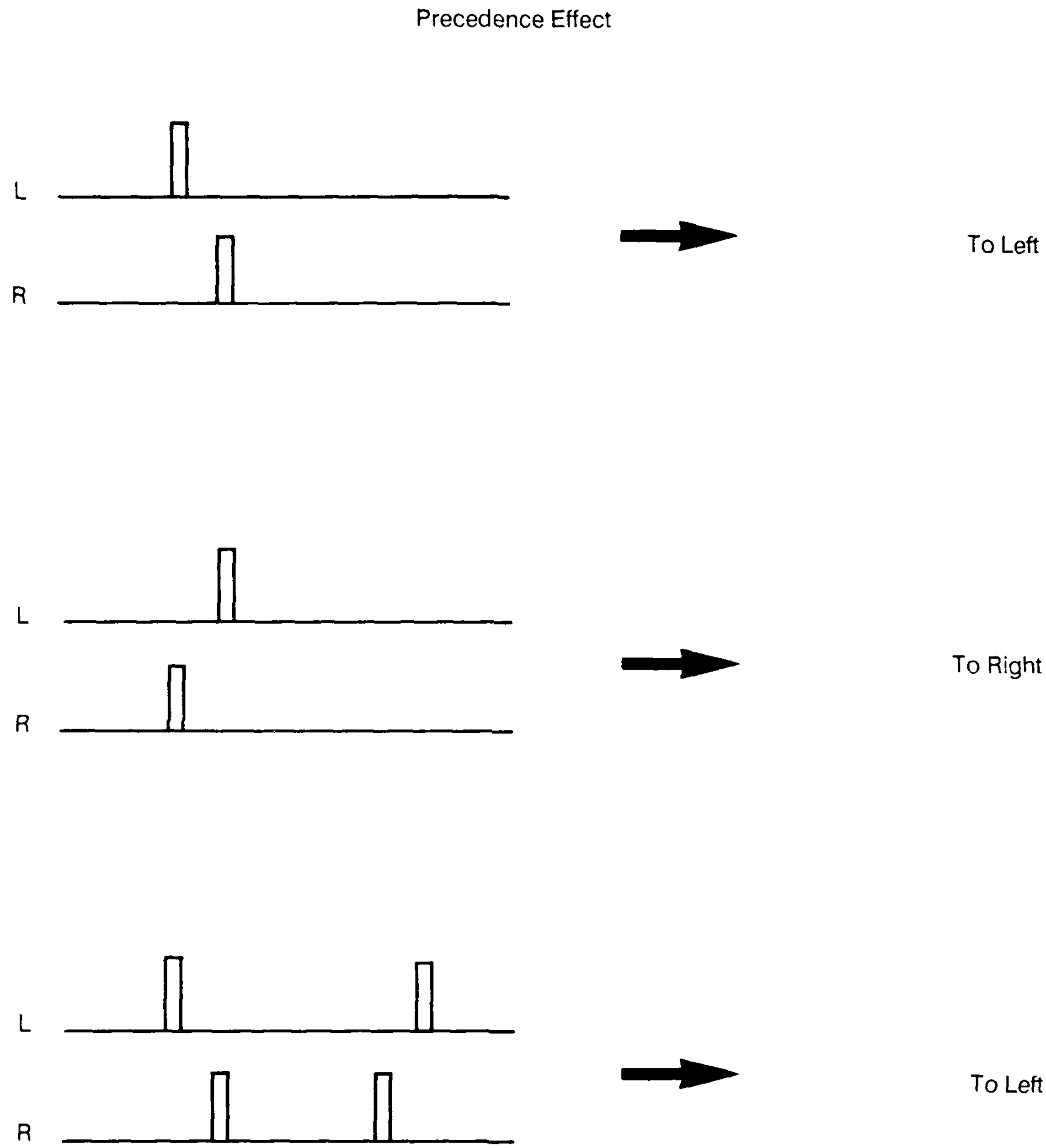


FIGURE 1 Precedence effect. Discrimination of pair two delay is poor (randomize pair one delay; response follows *on*). Preliminary indications are that the same result obtains with pair one and pair two separated in frequency.

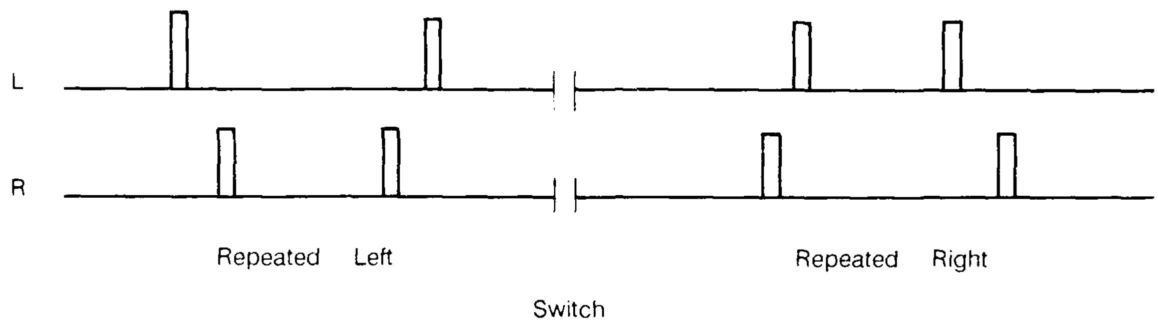


FIGURE 2 Clifton effect. Before the switch, a unitary left image is perceived; just after the switch, both the left and right images are perceived. Loss of precedence effect. Eventually (seconds later), a unitary right image is perceived.

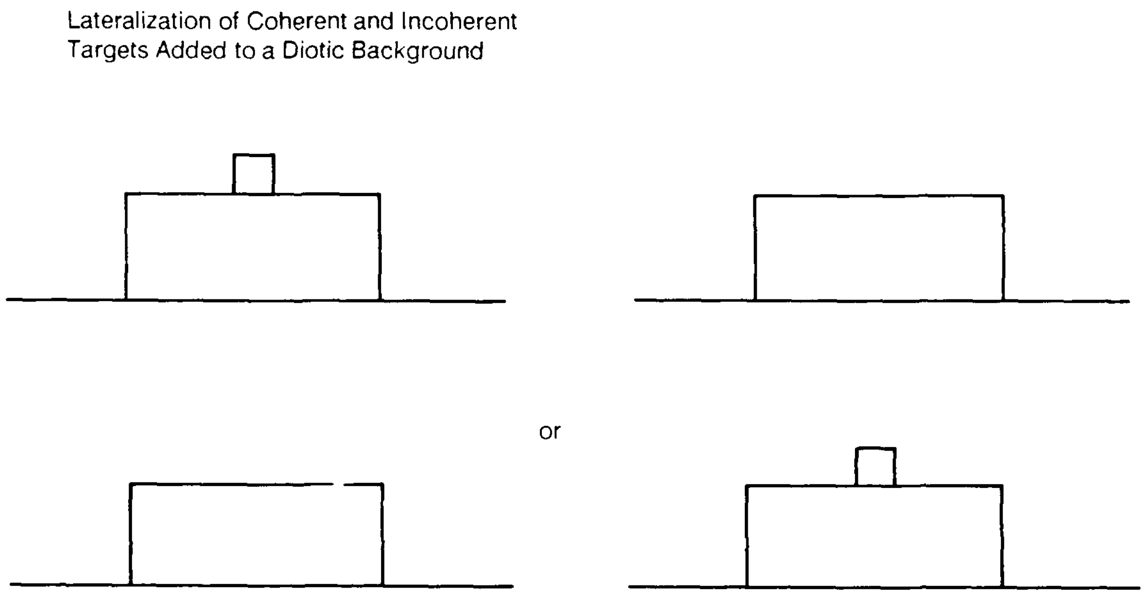


FIGURE 3 Schneider effect. (1) One can detect the presence of an incoherent target well before it is lateralizable (Egan and Benson). (2) Lateralization performance with incoherent increment is better than that with coherent increment of same average power. (3) Interaural difference distributions suggest that the coherent should give better performance since variability of interaural differences is less for this case.

ONGOING ENVELOPE DIFFERENCES

1. Lateral position is sensitive to interaural time delay in high-frequency waveforms with low-frequency envelopes. Examples include sinusoidally amplitude-modulated (SAM) tones and narrowband noise
2. SAM tones and noise band experiments show that lateral position is sensitive to ongoing envelope delay at all frequencies, as long as envelope modulation is maintained after internal filtering and the modulation rates are adequate

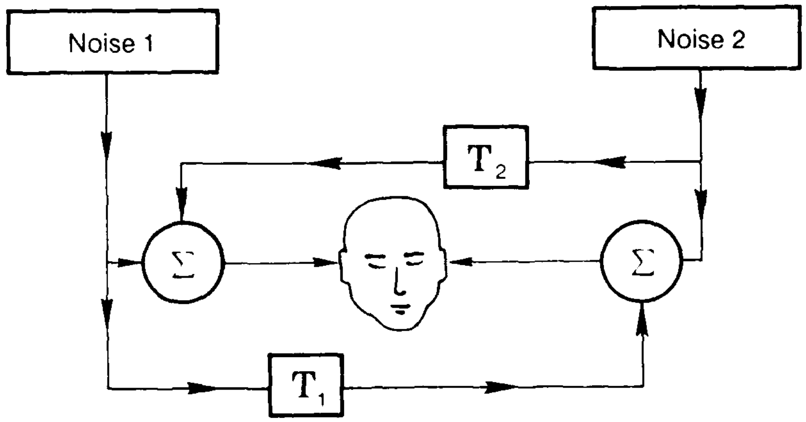


FIGURE 4 For steady, wideband noises, perceive broad, diffuse image. If one noise is switched off and then on, perceive two separate noise images for several seconds; after several seconds, it diffuses back to a single broad image.

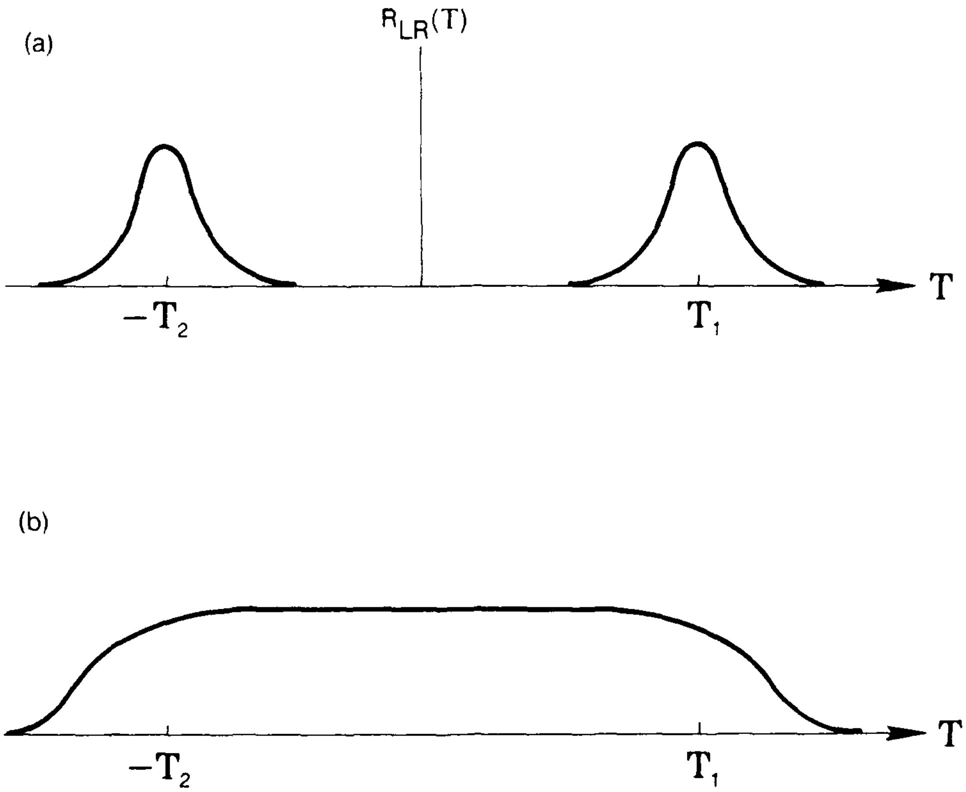


FIGURE 5 For two-noise stimulus, the wideband cross-correlation function is shown in (a) and the distribution of internal samples of interaural time delay (ongoing) in low-frequency filter outputs is shown in (b).

3. For a given delay and the same bandwidth, noise bands displace the image farther than SAM tones

LATERAL POSITION FOR WIDEBAND, CONTINUOUS STIMULI

1. A unitary, compact image is perceived when a fixed time delay or level difference is applied to a wideband noise waveform
2. Secondary images are perceived when anomalies in interaural differences are present in restricted spectral or temporal regions
3. Reductions in the interaural correlation reduces the compactness of the image

Time scales, time constants, and frequency ranges in lateralization are described as follows:

- Head width, ~1 ms
- Interaural phase sensitivity, below ~1,500 Hz
- Interaural time resolution, ~10 μs
- Internal delay line, ~1 ms
- Internal sampling time, ~ 10 ms. Filter impulse responses
- Output tracking rate, ~200 ms
- Cognitive reinterpretations ~2 s
- Long-term recalibration 2 h to (days?)

REFERENCE

Hafter, E.R., T.N. Buell, and V.M. Richards
1988 Onset-coding in lateralization: Its form, site and function. In G.M. Edelman, W.E. Gall, and W.M. Cowan, eds., *Functions of the Auditory System*. New York: Wiley.

Comment: Lateralization

ERVIN HAFTER

As discussants, we were asked to respond to the review papers to ferret out weaknesses, to point out differences of opinion, and to emphasize highlights. Fortunately, Colburn's excellent review was so thorough that the first two goals are unnecessary. Thus, I will focus my comments on some general points that arose repeatedly throughout the symposium, issues that are germane to the current status of the field as a whole.

In summarizing this symposium, one could say that the pervasive interest at this moment is in the ability to localize and identify complex sounds in natural environments, with an eye to an understanding of what constitutes an auditory object. Indeed, Green's admonition at the beginning of the symposium to "burn your earphones" and concentrate on "real" localization clearly depicts this trend. Nevertheless, while it is hard to deny that these are extremely interesting topics, it is clear from Colburn's review that most of the fundamental knowledge that we have about binaural hearing—the stuff from which theories are built—has come from experiments using headphones.

Headphones allow for control of the relevant binaural variables in ways not possible in the free field, and so it seems doubtful that their usefulness is over. Indeed, the argument that I present about binaural precedence and echo suppression is based primarily on data gathered using them. In his paper Colburn speaks of a group of studies by my colleagues and me in which the ability to detect interaural differences in trains of high-frequency clicks has been measured as a function of the rate at which the clicks are presented. The information transmitted by each click in the train can be found by comparing performance with trains of different lengths, and the primary result has been that lateralization relies most heavily upon interaural information in the signal's onset; we have called the process *binaural adaptation*. A major thrust of the work has been to discover just what it takes to terminate the adaptation so that the system can resample information in the binaural stimulus. Results show that such restarting can be produced by presentation of an additional, triggering signal; a list of effective triggers includes a brief moment of quiet inserted into the stimulus, brief bursts of noise, and short tones even of frequencies quite removed from those of the binaural stimulus. It would seem that this triggered resampling brings into question models that embody binaural precedence in the peripheral processes of binaural interaction. Figure 1 illustrates the point. The argument offered in the caption says that in the real-world situation of the cocktail party phenomenon, there are apt to be triggering signals that cause the binaural system to resample its environment. If so, should not there be misperceptions of direction when the disadapted system samples summations of stimuli arriving along separate pathways? One answer is that precedence represents, at least in part, a top-down, more cognitive solution to the problem of separating object from echoes. It says that precedence may be more akin to cases of perceptual dominance such as ventriloquism than to the midbrain processes of lateralization.

Expanding on the idea that more attention needs to be payed to cognitive processes, I would argue that much of our efforts in the future will show that what has been thought of as sound localization has influential top-down components. An example in point is the interesting phenomenon found by Rachel Clifton that has been discussed several times at this symposium. When her listeners were presented with trains of clicks that led slightly to the right loudspeaker, they heard the train as coming from the right, just as predicted by the precedence. However, when the order of presentation was reversed, with the left speaker

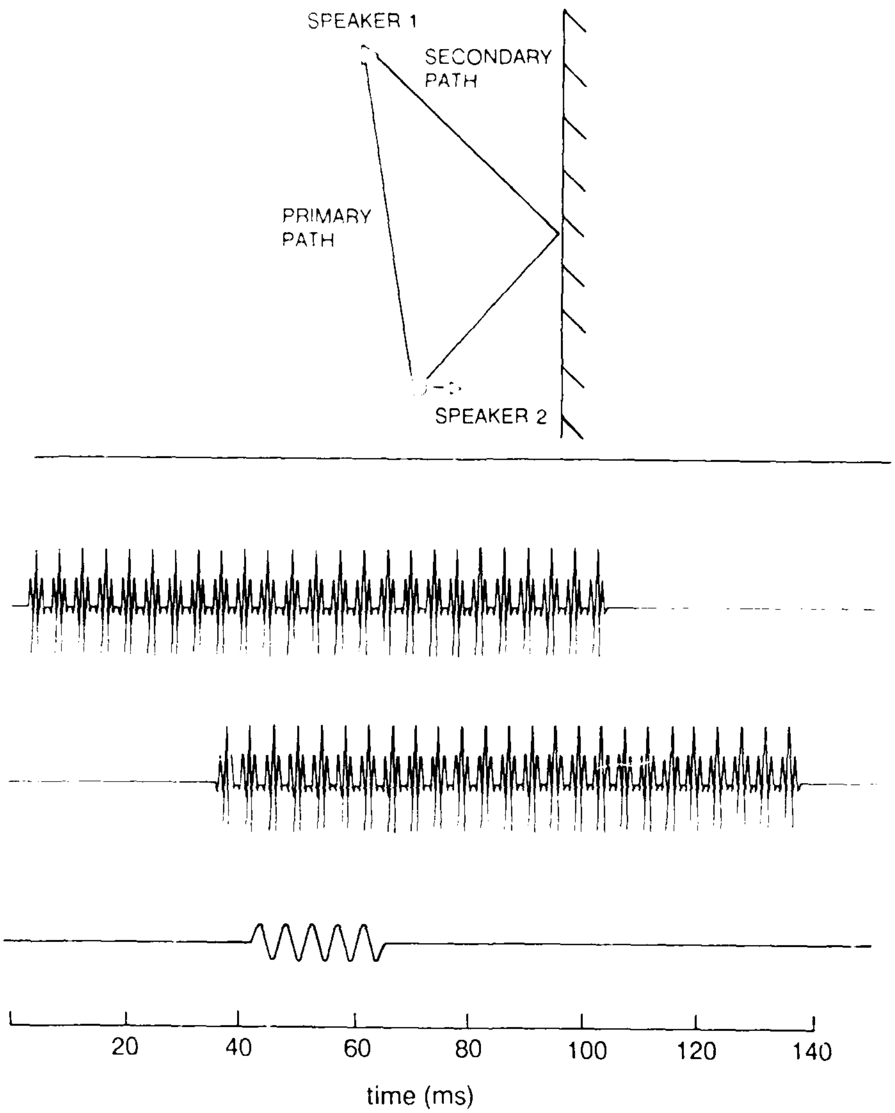


FIGURE 1 A hypothetical real-world situation in which one might expect a release from binaural adaptation. A 100-ms AM signal with a period of 4 ms is heard from a loudspeaker by a direct pathway and as a potential echo conducted along a path that is 10 m longer. These are shown in the two upper time lines. The short envelope period ensures binaural adaptation. After the arrival of the reflected version, the sounds at the two ears consist of the vector sums of the first and second wavefronts. For precedence to work, the listener must ignore interaural information in these sums. Suppose that a new sound is introduced into the environment (a tone pip?) from speaker 2. The question then is, "If this tone triggers resampling of interaural information, why does not the listener use the sums of first and second wavefronts to localize the modulated signal somewhere between speaker 1 and the echoic surface?" (Hafter, Buell, and Richards, 1988).

now leading, the summed image did not leap immediately to the left. Instead, the first few clicks after the reversal were heard as coming from both speakers. I view this as evidence of a cognitive process in which the auditory system has been asked to deal with an illogical situation in which a seeming echo suddenly precedes the primary signal. It is reminiscent

of Anne Treisman's famous demonstration in which a person is asked to "shadow" (repeat) a message presented only to the right earphone while ignoring the message to his or her left. When questioned, the shadower claims to know nothing of the unattended message; but, following a reversal of the wires to the earphones, the listener occasionally continues to shadow the old message for a while before switching over to the new right-ear message.

Both of these demonstrations seem to illustrate Hartmann's "plausibility hypothesis" (also discussed in this volume). These demonstrations show that perceptual processes higher than those generally associated with the hard-wiring of sound localization may have a strong influence on how the listener interprets the binaural data. My guess is that with the increased interest in more worldly stimuli in the free field, we will see more instances of cognitive/integrative aspects of spatial hearing.

REFERENCE

Hafter, E.R., T.N. Buell, and V.M. Richards
1988 Onset-coding in lateralization: Its form, site and function. In M. Edelman, W.E. Gall, and W.M. Cowan, eds., *Functions of the Auditory System*. New York: Wiley.

Neural Representations of Sound Location

JOHN C. MIDDLEBROOKS

The location of a sound source is not mapped directly onto a peripheral sensory sturcture, but it must be computed within the central nervous system from acoustical cues present at the two ears. Frequency-specific temporal and intensive cues to sound source location are provided by the interaction of the sound wave with the torso, head, and external ears. We know a great deal about how auditory information is analyzed in the frequency domain at the auditory periphery and about how some classes of spatial cues are extracted within the auditory brainstem. We know much less about how the multiple discrete foci of neural activity representing individual spatial *cues* might be interpreted at a higher level to represent a particular *location*. What is the pattern of neural activity underlying the image of a sound source localized in space? Are there central auditory structures within which a finite sound source elicits a restricted focus of neural activity? Several recent studies have used free-field stimulation to characterize the selectivity of single neurons for sound location. Two structures of the midbrain, the inferior and superior colliculi, provide examples of neural representations of auditory spatial cues and of sound locations. Results from studies at the midbrain level lead us to speculate on how the locations of sounds might be represented in the cerebral cortex.

The central nucleus of the inferior colliculus (ICC) is an obligatory link in the primary auditory path from the brainstem to the auditory forebrain. From studies using dichotic stimulation, the ICC has long been known to contain neurons that are sharply tuned for frequency and that are sensitive to acoustical cues for sound location. Recent studies using free-field stimulation have confirmed that the spatial tuning of ICC neurons can be accounted for by their sensitivity to particular binaural cues within restricted bands of frequency. In the ICC, spatial cues appear to be represented within discrete frequency-specific channels.

The superior colliculus is a sensorimotor integrative structure that generates orienting movements of the eyes, external ears, and head in response to input from a variety of sensory modalities. Neurons in the superior colliculus are broadly tuned for sound frequency. This broad tuning presumably reflects a convergence of spatial information from across the frequency dimension. When tested with free-field sound stimuli, neurons exhibit spatial tuning for the horizontal and vertical location of the sound source. A sound source produces a restricted focus of maximal neural activity that systematically shifts its position in the superior colliculus in response to changes in the location of the sound source. Thus, the superior colliculus appears to interpret multiple spatial cues to derive actual sound source *locations*.

Taking these midbrain structures as models. We can now address the issue of how auditory space might be represented in the cerebral cortex. The auditory cortex appears to be essential for sound localization, inasmuch as lesions there severely disrupt sound localization behavior in man and other mammals. Nevertheless, a cortical map of auditory space has yet to be identified. Spatial tuning has been studied systematically only in the primary auditory area (A1). Neurons in A1 resemble those in the ICC in that they are sharply tuned for frequency and are sensitive to binaural cues. About a third of neurons are selective for sound location, but as in the ICC, the spatial tuning of any given neuron seems to reflect sensitivity to an individual spatial cue, not a convergence of multiple cues representing a single location. One can infer that the spatial cues provided by a broad-band stimulus presented in the free field would elicit multiple discrete foci of activity in A1.

In contrast to A1, the largely unexplored second auditory area (A2) shares several basic auditory response properties with the superior colliculus. Neurons in A2 are broadly tuned for frequency and respond more reliably to broad-band stimuli than to tones, indicating a convergence of spatial information from across the frequency dimension. There are no anatomical connections from the superior colliculus to A2, but the ascending input to A2 seems to originate among some of the midbrain structures that provide auditory input to the superior colliculus. One might argue that the auditory map in the superior colliculus is not an appropriate model for a cortical representation, since the organization of the superior colliculus might reflect its close interactions with the motor system. However, the existence of an auditory map in the superior colliculus does serve as evidence that the nervous system is capable of deriving such a map. This, taken with the finding that spatial attributes of other sensory modalities are commonly mapped in the cortex, argues that a map of auditory space is likely to be present in the cortex. The parallels between the superior colliculus and A2 in their basic response properties and their sources of input make A2 an intriguing candidate for a cortical locus of auditory spatial representation.

Comment: Neural Representations of Sound Location

SHIGEYUKI KUWADA

Middlebrooks suggests that the superior colliculus is a site where the location of a sound is derived from the convergence of binaural cues from many frequency bands. His data suggest that, in the superior colliculus, convergence of inputs across frequency bands is primarily limited to the processing of interaural level differences (ILDs). However, there is another cue used to determine the azimuth of a sound source, interaural time differences (ITDs). Depending on their frequency tuning, binaural neurons can show ITD sensitivity to low-frequency sounds or to complex high-frequency sounds. Do ITD cues also converge in the superior colliculus?

A major function of the superior colliculus is to direct movements of the eyes, head, and ears in response to sensory stimuli of several modalities. It would be odd if ITDs were processed elsewhere. Although some studies on the superior colliculus have reported the presence of neurons that are sensitive to ITDs, the number of such cells is small compared with that for neurons sensitive to ILDs. The dearth of ITD-sensitive neurons in the superior colliculus is puzzling since a substantial part of the inferior colliculus is devoted to ITD processing. One possible explanation is that neurons sensitive to ITDs are present in the superior colliculus, but their existence is difficult to demonstrate in anesthetized

preparations. The data of Middlebrooks and others are consistent with this possibility. The neurons they study show a low response rate to sound. Furthermore, these neurons respond only transiently. By contrast, in an awake and performing monkey, neurons in the superior colliculus show a vigorous and sustained discharge to acoustic stimulation. If anesthesia has a slightly more potent effect on ITD-sensitive neurons than it does on ILD-sensitive neurons, then the presence of ITD-sensitive neurons would be difficult to detect.

Our laboratory has studied the ITD sensitivity of the inferior colliculus and medial geniculate neurons to tonal and complex stimuli, as well as the effects of anesthesia on ITD sensitivity. We discuss this material in relation to Middlebrooks' findings.

Models of Binaural Perception

RICHARD M. STERN

The last 10 years has seen major changes in the ways in which researchers in auditory perception have come to view the binaural system. Most models of binaural processing now examine the data in terms of the cross-correlation of the signals arriving at the two ears, after a stage of peripheral processing that includes band-pass filtering, rectification, and envelope detection at the higher frequencies. Consideration of the resulting patterns of assumed neural activity as a joint function of the center frequency of the peripheral band-pass filters and the delay parameter of the cross-correlation operation has proved to be useful in describing and predicting a wide variety of observed binaural phenomena presented through headphones. These models have also become increasingly sophisticated, thanks to advances in computational resources, new analytical techniques, and new insights into the phenomena themselves provided by recent experimental results.

Nevertheless, the application of formal quantitative and predictive theories to out-of-head localization phenomena has been extremely limited to date. This is a consequence of both the relatively small amount of experimental data on auditory localization available until recently and the complex and multidimensional nature of the stimuli and the resulting phenomena. Here we review and illustrate with examples some of the ways in which current binaural models based on interaural cross-correlation have been traditionally applied to binaural phenomena, discuss the application of the models to more recent data, and discuss some of the issues to be considered in extending the modeling process to out-of-head localization phenomena.

STATIONARY BINAURAL PHENOMENA

We first consider the lateralization of simple and complex stimuli with static interaural time and intensity differences. Several types of traditional models based on interaural cross-correlation can be used to predict the lateral position of low-frequency tonal stimuli (e.g., Stern and Colburn, 1978). These models extract position estimates in several ways, including the computation of the centroid with respect to the interaural delay parameter of the interaural cross-correlation of the stimulus (Figure 1). Recent studies have focused attention on more complex signals such as amplitude-modulated tones and band-pass noise. We have recently found that models based on the cross-correlation of peripheral auditory nerve activity can describe most aspects of the lateralization of these signals as well at both high and low frequencies. We also discuss the potential role that consistency of the

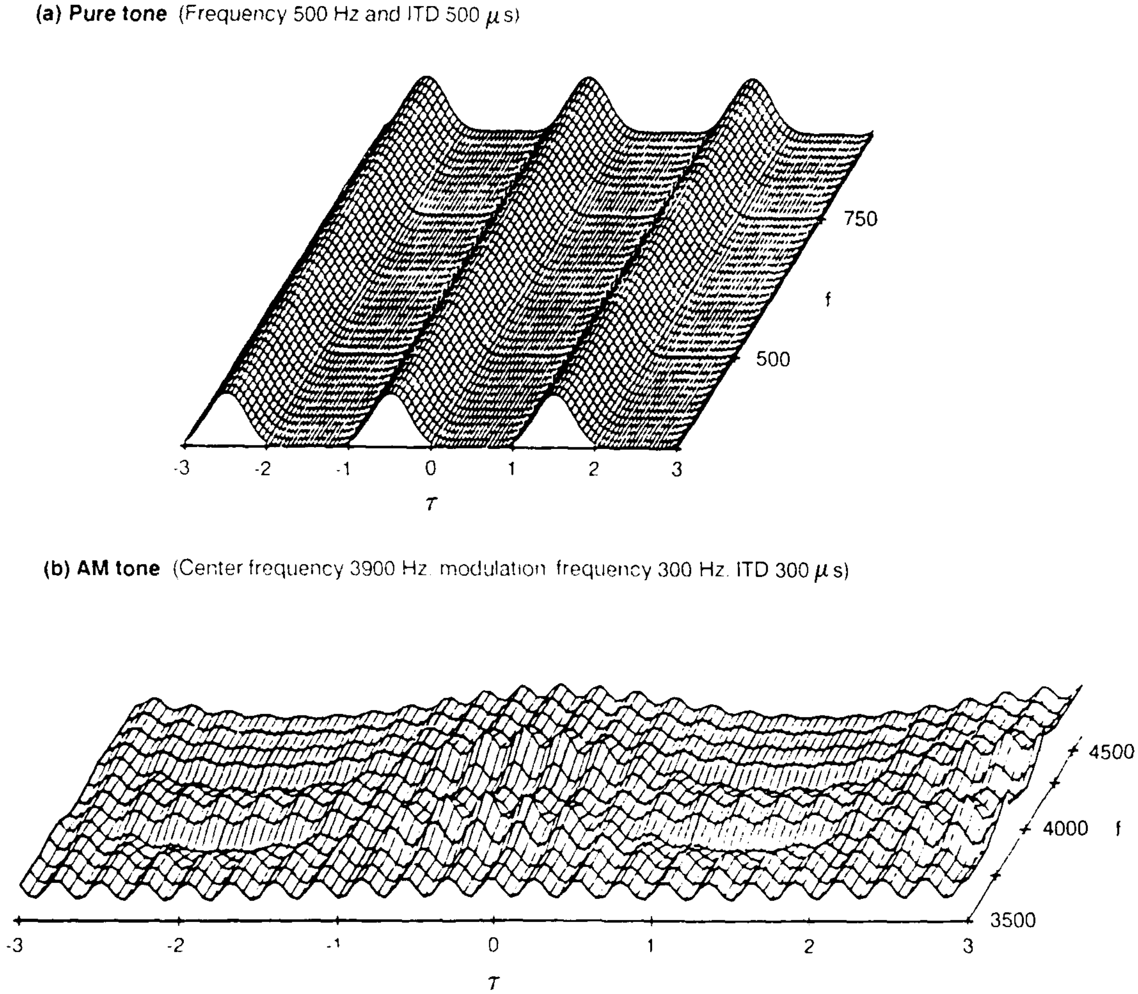


FIGURE 1 Examples of cross-correlation functions of typical stimuli used in lateralization experiments after processing by the auditory periphery. The functions are plotted as a joint function of the τ, the interaural delay parameter of the cross-correlation operation, and f, the center frequency of the filters used to model the peripheral auditory processing. (a) A 500-Hz tone presented with an interaural time delay of 500 μs. (b) An amplitude-modulated tone with a carrier frequency of 3,900 Hz, a modulation frequency of 300 Hz, and a waveform interaural time delay of 300 μs.

cross-correlation patterns over frequency may play in the lateralization of complex stimuli, in comparison with the relative salience of low-frequency envelopes versus the ongoing fine structure in the signals, as discussed by Stern, Zeiberg, and Trahiotis (1988).

In addition, we review the ways in which examination of patterns of interaural cross-correlation can be used to predict two related sets of phenomena: binaural masking-level differences and dichotic pitch phenomena. The presence of a target in a masking-level-difference experiment can be detected by noting the presence of decrements in cross-correlation in a local frequency region (Figure 2a). Models based on this principle are able to describe almost all of the classical literature on binaural masking-level differences (cf. Colburn, 1977). Similarly, many observed dichotic pitch phenomena can be predicted in

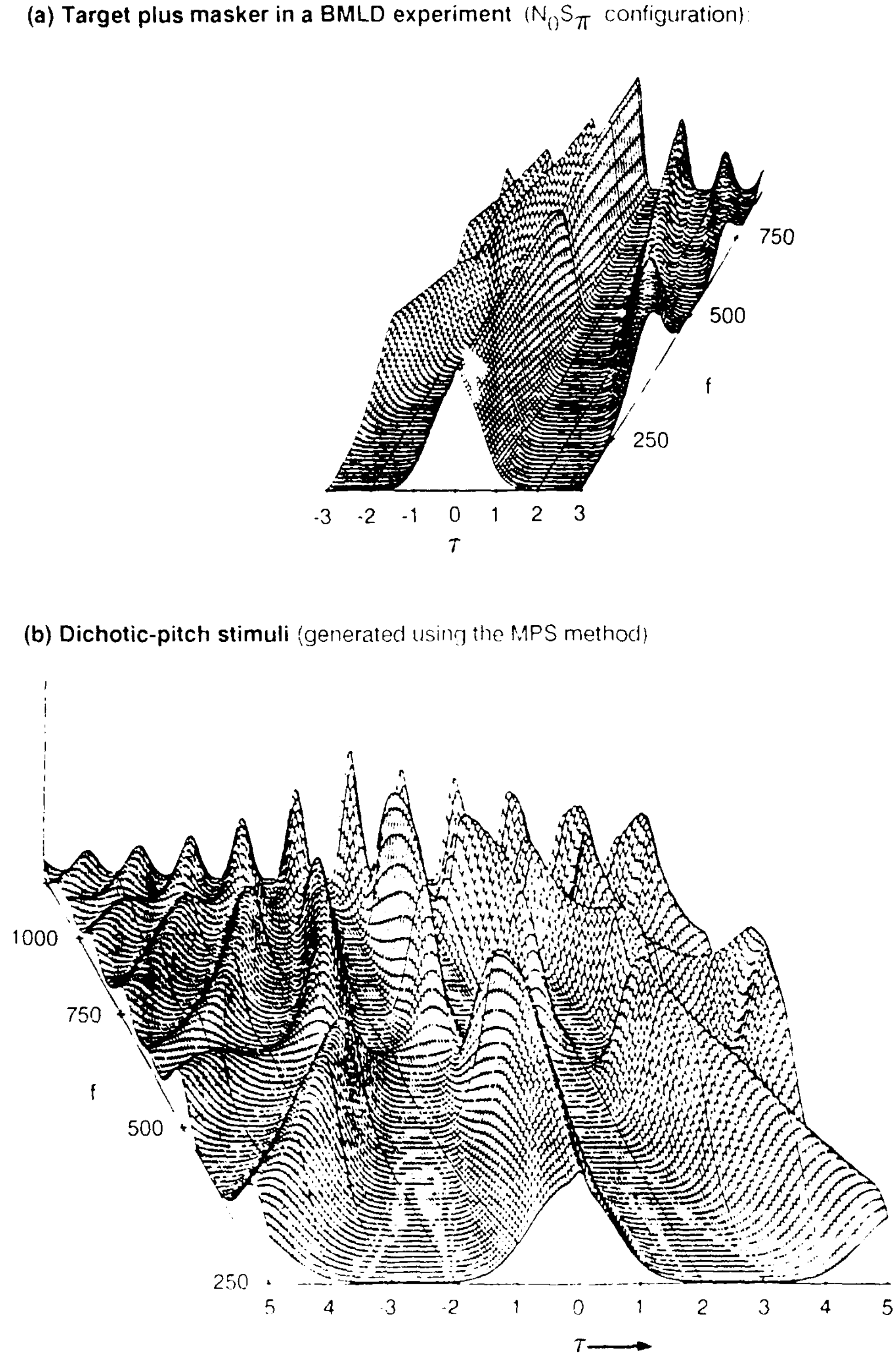


FIGURE 2 Examples of cross-correlation functions of stimuli used in other types of experiments. (a) Stimulus used in an $NoS\pi$ binaural masking-level difference experiment, with a target frequency of 500 Hz and a broadband noise masker. Note the decrease in correlation at frequencies near the target frequency. (b) Stimulus used in a typical dichotic-pitch experiment. This stimulus was generated using the multiple-phase shift (MPS) method.

terms of local discontinuities of the interaural cross-correlation function with respect to frequency (Bilsen, 1977) as illustrated in Figure 2b.

TIME-VARYING BINAURAL PHENOMENA

Several recent studies have addressed the perception of stimuli presented with a time-varying interaural time delay, intensity difference, or cross-correlation. The results of many of these experiments can easily be interpreted in terms of the temporal integration of the cross-correlation analysis of the binaural system.

Additionally, there has been an increased interest in the precedence effect and other short-term temporal phenomena that are commonly encountered in natural listening environments. We review the ways in which cross-correlation-based models have been extended to describe some of these results (e.g., Lindemann, 1986), as well as some of the problems encountered in attempting to construct practical signal-enhancement systems based on this type of analysis.

EXTENSIONS TO AUDITORY LOCALIZATION

Most models of actual localization phenomena have primarily considered the relative salience of different types of physical cues in the localization process, or they have predicted subjective aural attributes such as spaciousness, image fusion, and general listener preference as a function of various attributes of the cross-correlation function. In recent years, however, new experimental techniques have been developed that can provide physical measurements, psychophysical measurements, and simulations of localization phenomena under much more controlled conditions than were previously possible. We expect that the development of more systematic data about auditory localization will provide a major impetus toward the development of new models. We briefly review some of the current models of localization phenomena, and we discuss some of the issues that are likely to become significant for new theories as they are developed.

REFERENCES

Bilsen, F.A.
1977 Pitch of noise signals: Evidence for a "central spectrum." *Journal of the Acoustical Society of America* 61:150-161.

Colburn, H.S.
1977 Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *Journal of the Acoustical Society of America* 61:525-533.

Lindemann, W.
1986 Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *Journal of the Acoustical Society of America* 80:1608-1622.

Stern, R.M., Jr., and H.S. Colburn
1978 Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position. *Journal of the Acoustical Society of America* 64:127-140.

Stern, R.M., Jr., A.S. Zeiberg, and C. Trahiotis
1988 Lateralization of complex binaural stimuli: A weighted image model. *Journal of the Acoustical Society of America* 84:156-165.

Yost, W.A., and E. Hafter
1987 *Lateralization in Directional Hearing,* W.A. Yost and G. Guervitch, eds. New York: Springer-Verlag.

Comment: Models of Binaural Perception

H. STEVEN COLBURN

General issues in the development of mathematical and/or computational models are briefly discussed with particular attention to models of localization. Since a primary motivation for modeling is to test our understanding of the phenomena being modeled, explicit quantitative predictions must be possible from a complete model; otherwise, tests are incomplete. The modeling process should force us to capture our understanding in equations or algorithms and to refine vague notions into explicit formulations. A second major reason for modeling is to determine the relations among different phenomena. This is related to the ability to predict a wide set of phenomena with a common set of assumptions and is reflected in one of the measures of the quality of a model: The larger the set of data that can be described and the smaller the set of parameters that can be adjusted within the model, the better the model. The third major reason for developing models is that they can assist in insight formation. One gains insight in the initial formulation of the model as the assumptions are made explicit. It is often even more instructive when a model fails to describe a phenomenon or when a particular parameter or assumption is seen to be more critical than expected. When the structure of a model is considered, one is often led to interesting experiments that may be surprising predictions of the model or that may result from consideration of what would happen if an assumption were modified. An intimate interplay of modeling and experimentation is critical for productive research.

In the auditory localization area, these general considerations lead to several conclusions. First, models should be formulated to be able to be applied to as wide a set of phenomena as possible, including both psychophysical and physiological data. The recent physiological data on binaural interaction have not been effectively incorporated into quantitative models that attempt to address the general representation of binaural information. Second, vague concepts such as auditory object need to be captured in mathematical or algorithmic form. Third, there are many temporal phenomena that are not currently understood and that seem to require complex assumptions about the combination of information over time and frequency. The complexity of the assumptions seems to be related to the ease with which one can interpret the information as coming from distinct objects. Examples of phenomena of this type include the precedence effect, the Clifton effect, the Schneider effect, the Hafter effect, and others. Finally, the recent realization that several classical formulations of binaural processing models, such as the lateralization model, are unable to predict results from frozen-noise detection experiments (for example, the work of Gilkey, Robinson, and Hanna) illustrates both a significant current problem in our understanding and the importance of applying models to a wide class of data.

Some Sensory and Postural Influences on the Spatial Localization of Sound

JAMES R. LACKNER

Studies of auditory localization often are concerned with how changes in the physical dimensions of acoustic stimuli influence the perception of auditory direction. Such studies are usually carried out with the head fixed in position while the time, frequency, and intensity characteristics of the auditory signal are varied. The contribution of these factors to sound localization during static auditory localization is now reasonably well understood.

Under natural conditions, however, a person is freely moving about and his or her head and trunk position may vary both with respect to each other and to external objects. As a consequence, the auditory cues at the ears from a stationary sound source may change continuously. In order for the listener to hear the sound in the same external place, he must relate his changing position in space to the changing auditory cues at his ears resulting from the movement of his head.

The concern of the present review is to describe the wide range of postural information that is utilized in the computation of auditory direction and to show that auditory stimulation can also influence the apparent orientation of the body. One consequence of this reciprocal interaction between auditory and postural information is that the perceived location of a sound is determined not only by the pattern of physical stimulation at the ears but also by the registered orientation of the head in relation to the trunk and the gravitational vertical. Whenever there is an error in the registration of ongoing body orientation relative to the external environment, auditory mislocalizations and apparent auditory motion of related magnitude and time course can result. As a consequence, identical patterns of arrival time and intensity cues at the ears can give rise to the perception of sounds in widely disparate spatial positions in relation to the head and body and to the external environment, depending on the perceived representation of the body.

Perception of body orientation is dependent on multiple sources of afferent and efferent information concerning the spatial configuration of the body and its relation to the surroundings. Such information allows us, under normal circumstances, to preserve an accurate distinction between those changes in sensory and motor activity contingent on self-motion and those contingent on motion of or within the environment. Stable maintenance of this distinction provides an essential background for the ongoing control of normal body movement and posture. The range of sensory and motor inputs that influence orientation and the intricate ways in which they interact is only beginning to be understood. I summarize some of the known interactions implicating the somatosensory, vestibular, proprioceptive, and auditory systems.

Comment: Sensory Integration

ROBERT B. WELCH

A great many studies have demonstrated a very substantial influence of vision over audition, as seen most dratically in the ventriloquism effect. On the basis of this and other research, it has frequently been claimed that when the two sensory modalities are placed in conflict with one another, seeing *always* dominates hearing. However, a consideration of *why* vision should be so influential in certain situations leads to the prediction that it will play an *inferior* role in other situations.

It may be argued that the strong degree to which vision biases audition during spatial localization is due to the fact that vision is more appropriate (i.e., more accurate, precise) than audition for this particular task. Based on this notion, which can be referred to as the *modality appropriateness* hypothesis, it is predicted that the latter should dominate the former.

Such a task is temporal rate. Clearly, hearing is a much more temporally acute modality than is vision, as seen, for example, in the fact that the auditory flutter fusion threshold is much lower (i.e., a faster rate is required to attain it) than the visual critical flicker fusion threshold. In several experiments by the author and his colleagues, subjects were exposed to a blinking light and a repetitive beep in an otherwise dark room. The temporal rates of the two stimuli ranged from 4 to 10 H and in a given pairing they differed from each other by approximately 2 Hz. The subject was exposed to a particular visual-auditory pairing and instructed to report either the perceived auditory rate or the perceived visual rate.

As predicted from the modality appropriateness hypotheses, the perceived rate of the visual stimulus was strongly biased by the auditory rate, reversing the traditional dominance of vision over audition. Thus, for example, when an auditory stimulus of 4 Hz was paired with a visual stimulus of 6 Hz, the perceived visual rate was close to 4 Hz. The fact that the auditory stimulus was also biased slightly, but reliably, by the visual stimulus refutes the claim that the dominated sensory modality is merely suppressed or ignored.

More generally, it can be concluded that the nature of the resolution of two conflicting sensory modalities is determined by the appropriateness of the respective modalities for the task at hand. For those tasks in which, in everyday situations, one modality is much superior to the other, the bias of the first is quite substantial; for tasks in which the two modalities are about equally matched, the resolution of intersensory conflict represents a compromise.

Cross-Spectrum Effects in Localization: Parameters of Localizing Multiple Sources

WILLIAM A. YOST AND RAYMOND H. DYE

We are often able to determine the spatial location of a number of simultaneously active sound sources in our environment. We do this despite the fact that each ear receives a sound field made up of the sounds from the multiple sources (Figure 1). When there are multiple sound sources, the interaural differences of level and arrival time between the two ears created by this sound field vary across the spectrum. In fact, recent measurements (by Kuhn) indicate that interaural level and temporal differences measured at the ears differ across the frequency spectrum even for a single broad-band sound source. Thus, the auditory system must be able to process the interaural differences for each resolvable spectral component in the sound field arriving at the ears in order to determine the location of the various possible sound sources. Presumably, the binaural system can segregate the various sound sources in a complex sound spectrum into their probable sources by determining which components have interaural values that are consistent with a particular spatial location. There are few data describing human's ability to locate sounds in multisource environments.

Recent lateralization and localization investigations have revealed a few interesting findings concerning binaural processing of interaural differences across the audio spectrum. When there are a small number of components, the spectral and temporal arrangements of the components determine the degree to which the various components interact in determining binaural performance. For instance, if the components are harmonically related, one or more tones spaced octaves from a target tone can significantly elevate the threshold for processing interaural level and temporal differences of the target tone (R.H. Dye in Yost and Hafter, 1987; Buell, 1988; see Figure 2). It is as if all of the tones are seen by the binaural system as coming from one source, even though the various tonal components have different interaural values. Far less interaction among components is obtained when the tones are not harmonically related, and the tones with different interaural parameters appear as if they originate from different locations (Buell, 1988). When the spectrum of the sound has a large number of components, changing the interaural parameters of a subset of the components almost always segregates these components into a different source than that associated with the components that did not have their interaural values changed (Yost, Harder, and Due, 1987). Experiments in which subjects were asked to judge the spatial location of actual sources demonstrate that acuity in determining the spatial separation between sources is worse when there are two concurrent sources than when only one source at a time is presented (Perrott, 1984). Both localization (Thurlow and Martens, 1962) and lateralization (Wakefield, 1988) tasks show that the location of a target sound is displaced

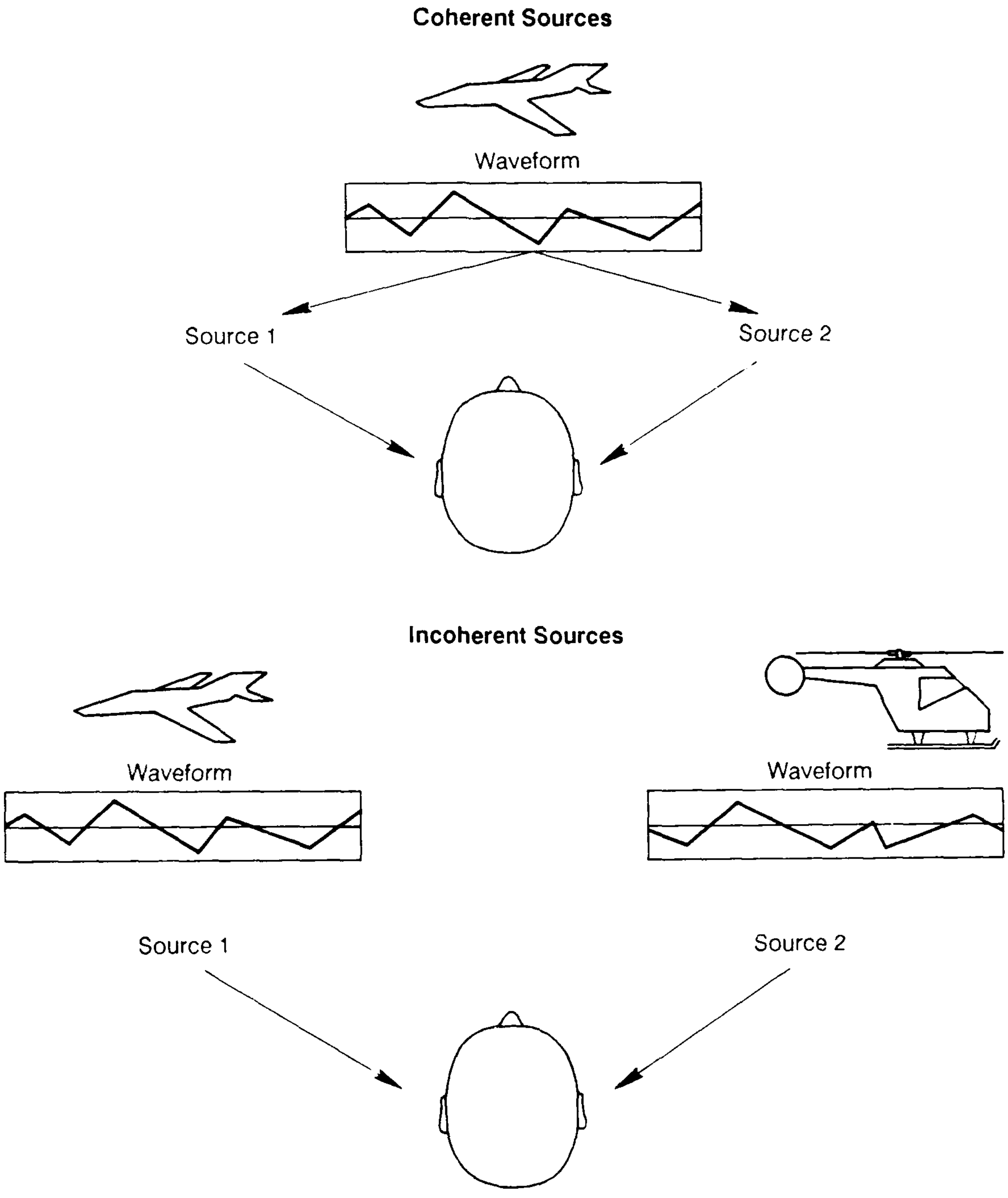


FIGURE 1 A depiction of the two major ways two or more sound sources have been presented to listeners. In the coherent condition both sources have the same waveform. This paradigm has been used to study a number of phenomena—precedence, coloration, and diffuse images—that occur in enclosed environments. Far fewer data are available for the incoherent conditions in which the sources contain different waveforms.

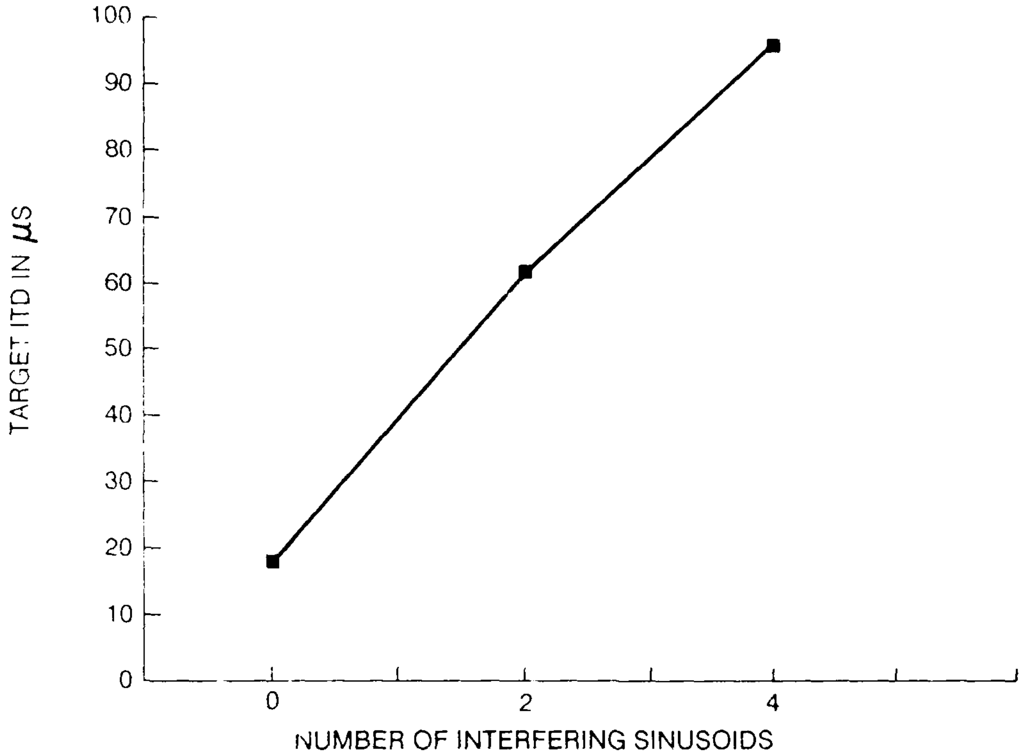


FIGURE 2 Data from Dye (in Yost and Hafter, 1987) showing the threshold for detecting an interaural time shift of a 750-Hz target tone in the presence of other simultaneously presented tonal components. For the zero number of interfering tones the listener discriminated a change in interaural time for the target alone; for two interfering tones the target was flanked by a 500- and 1,000-Hz tone; and for four interfering tones the target was flanked by 250-, 500-, 1,000-, and 1,250-Hz tones. As more tones are added to the complex, the sensitivity to the interaural temporal difference (ITD) of the target increases, even though the interfering tones are far outside the critical band of the target tone.

when another source is presented simultaneously. The overall conclusion from the limited data on localizing sounds in multisource environments is that listeners do not perform as well when more than one source is present as they do when only one source is present.

(This research was supported by the National Institute of Neurology, Communicative Disorders, and Stroke.)

REFERENCES

Buell, T.N.
1988 Ph.D. thesis. University of California, Berkeley.

Perrott, D.R.
1984 Concurrent minimum audible angle: A reexamination of the concept of auditory spatial acuity. *Journal of the Acoustical Society of America* 75:1201-1206.

Thurlow, W.R., and A.E. Martens
1962 Perception of steady and intermittent sounds with alternating noise-burst stimuli. *Journal of the Acoustical Society of America* 34:1853-1857.

Wakefield, G.
1988 Ph.D. thesis. University of Minnesota.

Yost, W.A., and E. Hafter
1987 *Lateralization in Directional Hearing*, W.A. Yost and G. Guervitch, eds. New York: Springer-Verlag.

Yost, W.A., P.J. Harder, and R.H. Dye
1987 Complex spectral patterns with interaural differences: Dichotic pitch and the central spectrum. In *Auditory Processing of Complex Sounds*, W.A. Yost and W.S. Watson, eds. New Jersey: Erlbaun Press.

Comment: Cross-Spectrum Effects in Localization

CONSTANTINE TRAHIOTIS

Comments were offered on the presentation of Yost and Dye as well as a consideration of other factors that influence listeners' ability to process interaural delays across spectral regions. New data concerning the detection of interaural delays in the presence of interfering noise were presented.

Auditory Motion Perception Via Successive "Snapshot" Analysis

D. WESLEY GRANTHAM

Evidence is presented suggesting that the processing of horizontal motion by the auditory system is accomplished by a snapshot analysis; that is, rather than being appreciated directly, the velocity of a moving auditory target is *inferred* by successive comparison of the target's different spatial positions at different instants in time. In all experiments horizontal movement of a 500-Hz tone was produced by dynamically varying the intensities of the inputs to two fixed loudspeakers in an anechoic chamber to produce moving phantom images according to Bauer's law of sines. In the first experiment, auditory spatial acuity was measured with both moving and stationary signals. For all conditions tested the dynamic measure of acuity (the minimum audible movement angle) was equal to or worse than the static acuity measure (the minimum audible angle), indicating that there is nothing special about movement in terms of spatial acuity. In the second experiment, increment thresholds were measured for the velocity of horizontally moving sounds. For a variety of conditions (including different stimulus durations and reference velocities), the relevant cue for performing this task turned out to be not velocity but angular distance traversed. Furthermore, when spatial cues were eliminated (by randomizing the durations of the two intervals in the two-interval forced choice task), velocity increment thresholds increased by more than a factor of 2, indicating again that velocity *per se* was apparently not the cue employed by subjects to perform this task.

The third experiment asked the following question: Is auditory velocity perceived directly, or is it a secondary characteristic derived from the prior discrimination of spatial and temporal positions? This problem was addressed by having subjects discriminate the velocity of various pairs of horizontally moving sounds. For a given pair the velocity difference could be produced by presenting moving sounds (1) whose durations were equal, but whose angular extents differed; (2) whose angular extents were equal, but whose durations differed; or (3) that differed both in angular extent *and* duration. A comparison of velocity discrimination for these pairs of stimuli revealed that velocity does *not* form a unique perceptual dimension in audition; the discrimination of velocity differences in stimuli that differ in both time and space is no better (and is often worse) than discrimination of velocity differences in stimuli that differ along only one of the two dimensions.

Thus, for the stimuli and range of conditions employed in these studies, the perception of auditory velocity appears to be *derived* from the prior discrimination of spatial and temporal differences.

Comment: Are There Motion Detectors in the Auditory System?

DAVID R. PERROT

If Grantham's snapshot hypothesis is correct, then it would appear that the auditory modality developed a different solution to the problem of motion than that taken by the visual system. The results of a number of experiments which seem to indicate that motion *can be directly appreciated* are discussed. For example, subjects can make accurate judgments regarding the velocity of the source even in situations in which other cues (e.g., duration and distance traveled) have been eliminated. Similarly, if the same temporal constraints are imposed, motion detection may be superior to the performance observed under static listening conditions.

These results could potentially be incorporated into Grantham's snapshot hypothesis, but it seems unlikely. In the absence of some crucial experiment, it would appear to be too early to throw out the notion of specialized neural elements tuned to the translocation of acoustic events.

Measurement and Interpretation of Head-Related Transfer Functions

FREDERIC L. WIGHTMAN, DORIS KISTLER, AND MARIANNE ARRUDA

The major acoustical cues for sound localization, interaural time and intensity differences and direction-dependent spectral shaping, are generated by the interaction of an incoming sound wave with a listener's head and pinnae. These acoustical cues, embedded in the head-related transfer function from a particular free-field position to a listener's eardrum, have been extensively studied; and their theoretical bases are reasonably well understood, as documented by previous reviews in this volume. In contrast, many questions remain unanswered about the coding and processing of the acoustical cues in the human auditory system. In recent years there has been a resurgence of research in this area, stimulated in part by the availability of sophisticated stimulus control techniques. One of these techniques involves simulating the naturally occurring acoustic localization cues by digital processing algorithms designed to mimic the acoustic effects of a listener's head and pinnae. The success of the simulation techniques depends on the accuracy with which a listener's head-related transfer functions are measured, since these transfer functions are the basis of all of the processing algorithms. Our presentation focuses on the practical issues involved in the measurement of head-related transfer functions and factors that affect the accuracy of the measurements. Such issues are signal-to-noise ratio, bandwidth, spectral and spatial resolution, sources of variability such as microphone position and stability, and the generalizability of measurements made from dummy heads. In addition, we present data on the dependence of the head-related transfer function on source azimuth and elevation, on the intersubject differences in these dependencies, and on how these intersubject differences relate to intersubject differences in sound localization behavior (Figure 1).

Table 1 shows summary data from the free-field and headphone listening conditions. In order to prepare this table, we computed the judgment centroid for each of the 72 source positions in both free-field and headphone conditions for each of eight subjects. We then computed correlations between target and centroid azimuth and between target and centroid elevation for both the free-field and headphone conditions. Finally, we computed a three-dimensional goodness of fit between the set of points defined by each subject's judgment centroids and the set of points defined by the target locations. The correlations and the goodness of fit were computed according to algorithms devised by Schonemann and colleagues for the purpose of fitting one matrix to another. Our use of the algorithms involves rigid rotation of the matrix of centroids to a least-squares fit with the matrix of target positions (thus, we ignore constant azimuth and/or elevation biases in the judgments) and computation of a statistic S, the normalized sum of the squared residuals. The final

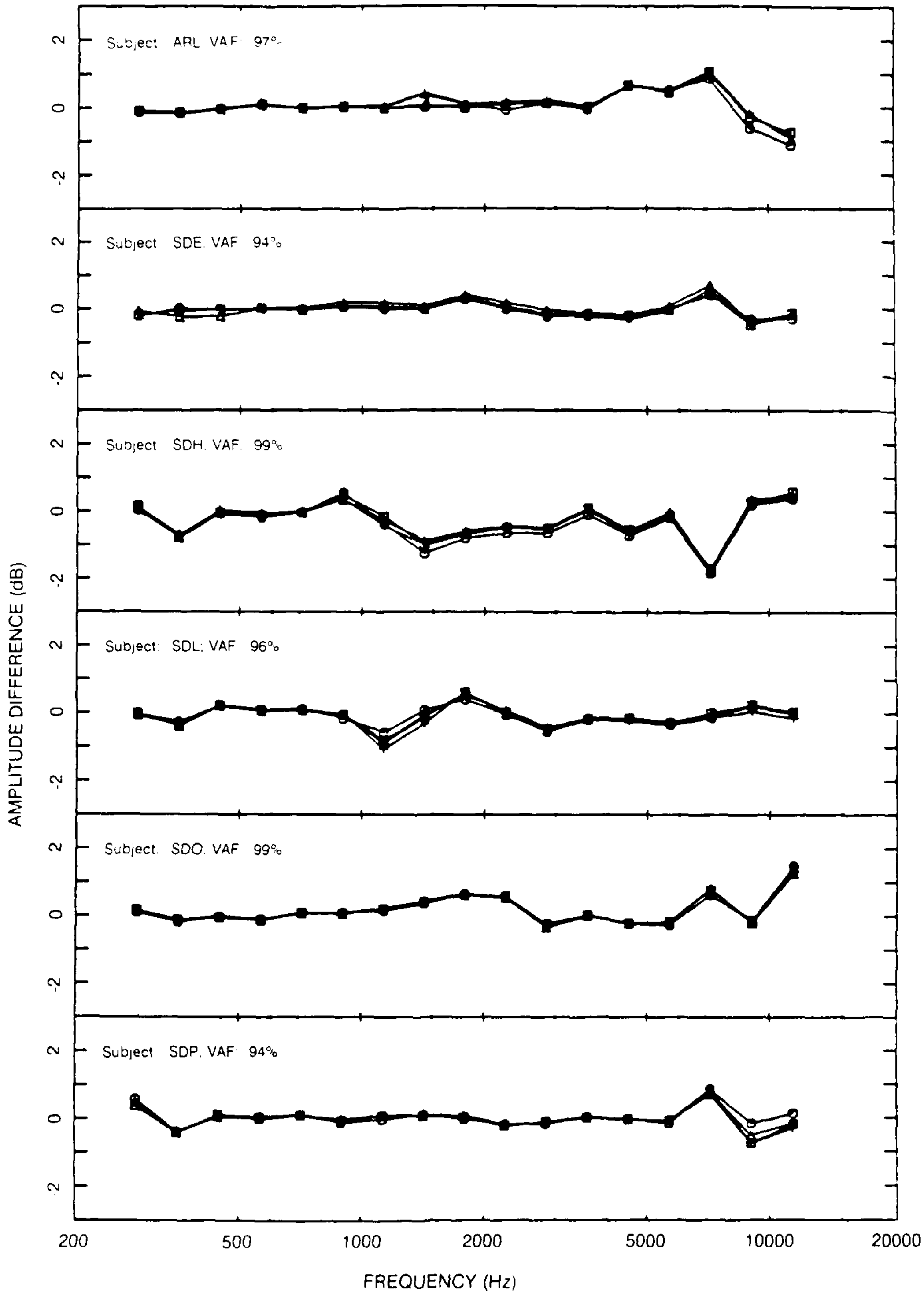


FIGURE 1 Results from the acoustical verification experiments. Data are shown from six subjects and four source positions each. The different symbols represent different source positions. The six panels show the 1/3-octave amplitude difference in decibels, measured at the listener's eardrum, between the spectrum of a free-field stimulus and the spectrum of a synthesized stimulus presented over headphones. The percentage variance accounted for (VAF) in each panel reflects the extent to which the pattern of differences as a function of frequency is constant across the four positions.

TABLE 1 Global Measures of Localization Performance

Identity	Goodness of Fit	Azimuth Correlation	Elevation Correlation	Percentage of Reversals
SDE	.93 (.89)	.983 (.973)	.68 (.43)	12 (20)
SDH	.95 (.95)	.965 (.950)	.92 (.83)	5 (13)
SDL	.97 (.95)	.982 (.976)	.89 (.85)	7 (14)
SDM	.98 (.98)	.985 (.985)	.94 (.93)	5 (9)
SDO	.96 (.96)	.987 (.986)	.94 (.92)	4 (11)
SDP	.99 (.98)	.994 (.990)	.96 (.88)	3 (6)
SED	.96 (.95)	.972 (.986)	.93 (.82)	4 (6)
SER	.96 (.97)	.986 (.990)	.96 (.94)	5 (8)

NOTE: Measures of free-field performance are followed by measures of simulation performance in parentheses.

measure, which we call correlation in one case and goodness of fit in another, is equal to 1 − S)$^{1/2}$. In the case of the azimuth or elevation correlations, this measure is nearly identical to a Pearson correlation since the data are two-dimensional. The goodness of fit measure is essentially a three-dimensional Pearson correlation, which gives an overall indication of the degree of match between the targets and the judgment centroids. The percent reversals entries represent the percentage of judgments that could clearly be classified as front-back reversals.

BIBLIOGRAPHY

Blauert, J.

1983 *Spatial Hearing: The Psychophysics of Human Sound Localization.* Cambridge, Mass.: MIT Press.

Butler, R.A.

1975 The influence of the external and middle ear on auditory discrimination. In *Handbook of Sensory Physiology*, W.D. Keidel and W.D. Neff, eds. Berlin: Springer-Verlag.

Mehrgardt, S., and V. Mellert

1977 Transformation characteristics of the external human ear. *Journal of the Acoustical Society of America* 61:1567-1576.

Oldfield, S., and S. Parker

1984 Acuity of sound localization: A topography of auditory space. I. Normal hearing condition. *Perception* 13:581-600.

Wightman, F.L., and D.J. Kistler

1989a Headphone simulation of free-field listening. I. Stimulus synthesis. *Journal of the Acoustical Society of America.* In press.

1989b Simulation of free-field listening with headphones. II. Psychophysical verification. *Journal of the Acoustical Society of America.* In press.

Comment: Measurement and Interpretation of Head-Related Transfer Functions

CHRISTOPH POESSELT AND JENS BLAUERT

It has been known for a long time—at least since the work of Batteau in the late 1960s—that authentic reproduction of auditory perspective can be achieved in the following way. Pick up the signals at the eardrum of a listener, record them, and play them back after careful equalization such that identical signals are presented to this listener's eardrums as have been present there during the pick-up session. Distortion of the auditory perspective can be traced down to distortions of the input signals to the eardrums.

We have conducted a number of psychoacoustic experiments in which we introduced characteristic linear distortions as can typically be seen in binaural transmission systems caused by insufficient equalization. Among other things, we report on the effect of interindividual averaging of head transfer functions, on the effects of truncating the external ear impulse responses, and on what happens when one listens through somebody else's external ears.

Part II
Applications of Localization Data and Theories

Application of Auditory Spatial Information in Virtual Display Systems

ELIZABETH M. WENZEL, FREDERIC L. WIGHTMAN, AND SCOTT H. FOSTER

With rapid advances in technology and the concomitant requirement for managing and interpreting complex systems of information, an increasing amount of applied research has been devoted to reconfigurable interfaces like the virtual display. Indeed, the kind of artificial reality once relegated to the specialized world of the cockpit simulator is now being seen as the next logical step in interface development for many types of advanced computing applications.

At the Ames Research Center of the National Aeronautics and Space Administration (NASA), a head-mounted, wide-angle, stereoscopic display system controlled by operator position, voice, and gesture has been developed to provide a multipurpose, uniform interface for users with different levels of skill and training in a variety of tasks (Fisher et al., 1988). This virtual interface environment workstation (VIEW) generates a multimodal, interactive display environment in which a user can virtually explore a 360-degree synthesized or remotely sensed world and viscerally interact with its components. Primary applications of the system include telerobotics/telepresence control, air traffic control displays, management of large-scale, integrated information systems such as those anticipated for the Space Station, and visualization of complex, multidimensional data in such diverse fields as computational fluid dynamics, surgical planning and simulation, modeling of biochemical molecular interactions, and mechanical and architectural design.

As with most research in information displays, virtual displays have generally emphasized visual information. Many investigators, however, have pointed out the importance of the auditory system as an information channel. A three-dimensional auditory display could take advantage of intrinsic sensory abilities like localization and perceptual organization by generating dynamic, multidimensional patterns of acoustic events that convey meaning about objects in the spatial world. Applications involve any context in which the user's situational awareness is critical, particularly when visual cues are limited or absent and work load is high. Such a display would generate localized cues in a flexible and dynamic manner. Whereas this can be readily achieved with an array of real sound sources or loudspeakers, the custom signal-processing board being developed at NASA-Ames maximizes flexibility and portability by synthetically generating three-dimensional sound in real time for delivery through headphones (Wenzel et al., 1988).

Psychoacoustic research suggests that perceptually veridical localization over headphones is possible if both the direction-dependent pinna cues and the more well understood cues of interaural time and intensity are adequately synthesized (Blauert, 1983). Although

the real-time prototype is not yet complete, recent studies by Dr. Wightman and his colleagues have confirmed the perceptual adequacy of the basic approach to synthesis utilizing measurements of head-related transform functions (HRTFs) (Wightman and Kistler, 1988a,b).

Research devoted to understanding the basic mechanisms and processes of human sound localization will have a critical impact on the general utility of a three-dimensional auditory display in any context. However, it should also be remembered that the application of such knowledge may impose its own constraints. The goal of a three-dimensional auditory display is to present unambiguous spatial information as flexibly, dynamically, and efficiently as possible, often under conditions that are less than ideal for detecting subtle acoustic cues. This is particularly true for the cockpit of a military jet or the helicopter during nap-of-the-earth flight, where the acoustic environment is extremely noisy, the sensitivity and bandwidth of the transducing equipment is limited, and the pilot often has noise-induced hearing loss, yet dependence on the auditory channel is unusually high because of the high visual and motor work load. On the other hand, the more generic virtual environments exemplified by the VIEW system will be less subject to such stringent requirements, offering applications in which the full potential of auditory cueing may be explored.

Ultimately, many factors will need to be considered and many compromises made in attempting to produce a veridical acoustic display. These can be loosely categorized as (1) practical signal-processing issues like the required bandwidth, frequency resolution, and computational precision for an adequate synthesis; (2) the role of individual differences and the possibility of overcoming such effects through adaptation to non-listener-specific transforms or even enhancement of features of the pinna cues to form a set of generalized HRTFs; (3) factors that promote externalization via interaction with the other senses, including correlated visual stimuli, dynamic cueing induced by source motion or head-coupled motion, and the creation of veridical acoustic spaces by modeling distance and reverberation effects; and (4) the determinants of perceptual organization in acoustic signals, that is, the characteristics of complex stimuli that aid localization accuracy and enhance discriminability of multiple sources. In the past it has not been possible to adequately test many aspects of these questions simply because it was technically too difficult to put the stimuli under direct experimental control. It is hoped that real-time signal-processing devices like the prototype under development at Ames will prove to be useful tools for examining some of these issues as well as for furnishing the basic technology for sophisticated acoustic displays.

REFERENCES

Blauert, J.

1983 *Spatial Hearing: The Psychophysics of Human Sound Localization.* Cambridge, Mass.: MIT Press.

Fisher, S.S., E.M. Wenzel, C. Coler, and M.W. McGreevy

1988 Virtual interface environment workstations. *Proceedings of the Human Factors Society* 32:91-95.

Wenzel, E.M., F.L. Wightman, and S.H. Foster

1988 A virtual display system for conveying three-dimensional acoustic information. *Proceedings of the Human Factors Society* 32:86-90.

Wightman, F.L., and D.J. Kistler

1988a Headphone simulation of free-field listening I: stimulus synthesis. *Journal of the Acoustical Society of America.* In press.

1988b Headphone simulation of free-field listening II: psychophysical validation. *Journal of the Acoustical Society of America.* In press.

A Real-Time Digital Auditory Localization Cue Synthesizer

RICHARD L. MCKINLEY AND MARK A. ERICSON

This presentation describes the design and performance of an auditory localization cue synthesizer that is coupled to head position. The data include salient parameters of the design and direct comparison of localization performance with the synthesizer to free-field performance in human subjects. The current data are for azimuth only. The synthesized stimuli are presented over headphones and for most listeners appear to be out of head and are easy to localize. The synthesizer uses a single audio input and is controlled via a standard RS-232 interface.

Auditory Psychomotor Coordination: Auditory Spatial Information Can Facilitate the Localization of Visual Targets

DAVID R. PERROTT

A series of experiments are described in which subjects were required to locate and identify a small (0.5 degree) visual target within a large (240 degree) test field. Visual search time was determined as a function of the position of the target relative to the initial position of the observers' eyes. Concurrent presentation of a 10-Hz click train from the same location as that occupied by the visual target substantially reduced the visual search time for all events within this extended field. The advantages provided by the presence of a spatially correlated sound were evident even when the visual target was located within 10 degrees of the initial line of gaze. The utility of auditory spatial information in visual target acquisition was most apparent when the position of the visual target was also free to vary in the vertical dimension (±46 degree). Nonspatial tonal cues were considerably less effective in reducing the visual search period than when spatially correlated sounds were available, even when the subjects were tested monaurally in the latter condition. While the notion that human subjects can utilize auditory spatial information to redirect the position of their eyes, head, and body is not new, for example, it is a familiar aspect of the orientation reflex, the current results indicate that auditory-directed movements are sufficiently precise to allow very rapid acquisition of a visual event. It is interesting to note that the relatively poor resolution of the auditory spatial system (seldom better than 1 degree) may be viewed, in this context, as entirely adequate to the demands of this spatial task. If a primary responsibility of the auditory system is to provide the spatial input to the motor system, resolution of only a few degrees would be adequate to bring the target within the central visual field.

Auditory Heads-Up Display: Observer Use of Cues Correlated with Head Movement

ROBERT D. SORKIN

Commercial and military pilots must process an immense amount of data about the flight environment. It has been suggested that pilot performance could be improved by providing (via headphones) a three-dimensional auditory display of information, for example, the azimuth and elevation of air and ground targets. The apparent position of these displayed signals, relative to the aircraft, would be invariant with respect to the orientation of the pilot's head. We were interested in testing some of the assumptions implicit in the suggestion: Does coupling head movement to a directional auditory stimulus improve localization of the stimulus? What are some performance effects of displaying a spatial stimulus that moves appropriately or inappropriately as the head moves?

An auditory heads-up display (AHUD) system was implemented by synthesizing an array of headphone signals designed to yield externalized percepts of a target at 96 different locations (16 different azimuths by 6 different elevations; the signals were provided by F. Wightman and D. Kistler of the University of Wisconsin). An observer was placed in a cockpit mock-up surrounded by a painted scene of horizon, ground, and sky. After listening to a sequence of signals, the observer reported the target's location. In the normal AHUD operating mode, information from a magnetic sensor (ISOTRAK) system was used to sense the position of the observer's head and correct the headphone stimulus so that the apparent position of the target was approximately fixed in space. Three different conditions relating the observer's head movement to the target's spatial position were studied: (1) the normal AHUD mode with the observer instructed to immediately face toward all targets; (2) a fixed mode with the observer instructed to maintain a straight-ahead, level view; and (3) an uncorrelated mode with the position sensor disabled and the observer told to face all targets (Figure 1). The resulting localization performance was limited by the coarseness of the stimulus array, the bandwidth of the headphone system, and the procedure for constructing the target signals. Accuracy of localization was best in the normal AHUD condition. This experiment shows that correlated auditory and nonauditory cues generated by voluntary movement of an observer's head can improve target localization.

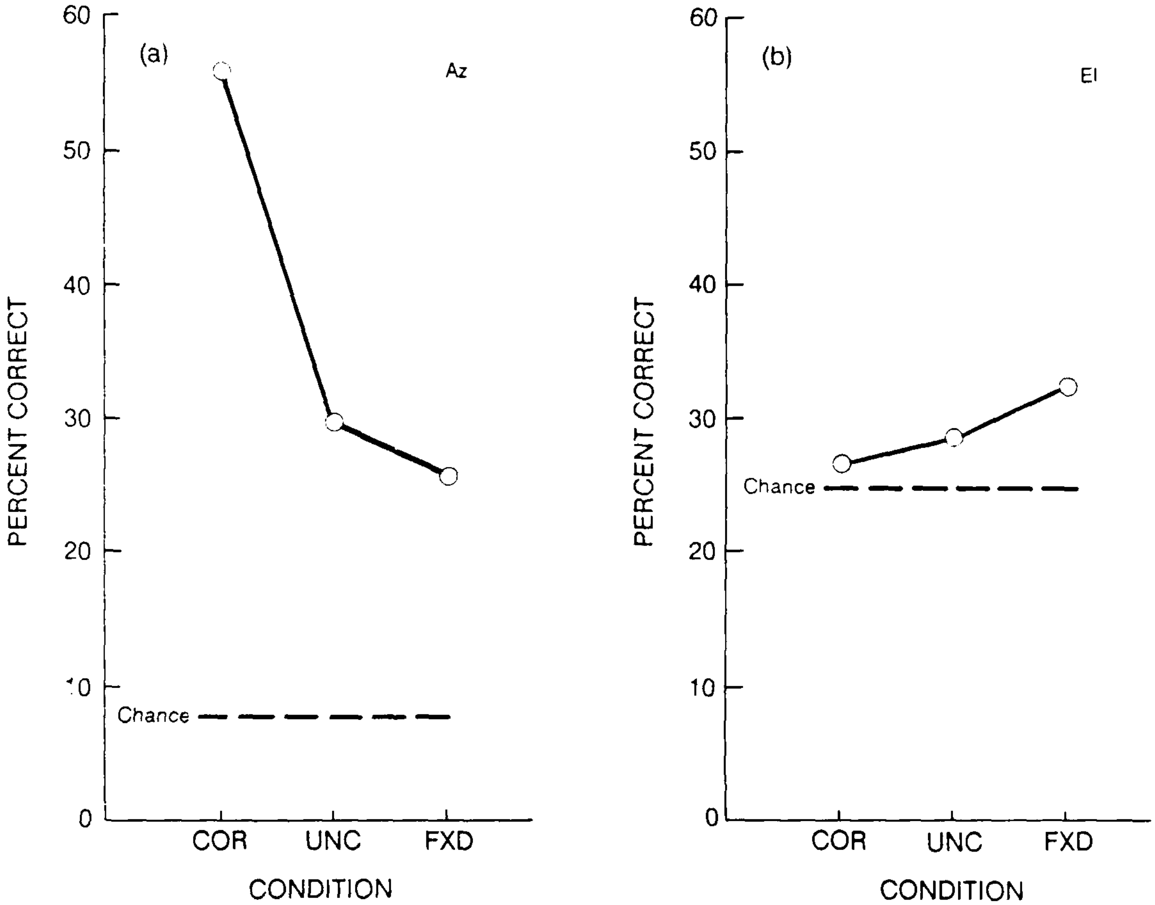


FIGURE 1 Accuracy of target localization in the azimuthal plane (a) and in the elevational plane (b) is above chance for all three modes: correlated, uncorrelated, and fixed.

Localization of Sound in Rooms

WILLIAM M. HARTMANN

When sounds are presented in a room, most of the energy arriving at the listener's ears is from sound waves that have been reflected from the surfaces of the room. Reflected sound waves seriously confound those acoustical cues that are known to be used by the listener in localizing sound, that is, spectral character and interaural differences in the arrival time, or the intensity, or the spectrum. The fact that we can localize sounds in a room at all is normally ascribed to the precedence effect, as studied by means of click pairs by Wallach, Newman, and Rosensweig. When one proceeds beyond the click-pair paradigm, a number of questions immediately arise. What is the relationship between the precedence effect, as observed in localization studies, and the reduction of perceived reflection and reverberation commonly called the Hass effect, as observed in studies? Can the rather different time scales for these two effects be comprehended in a single model? How should one think about the precedence effect when the duration of the stimulus is longer than the reflection delay times of the room? How does one explain the fact that broad-band noise and complex tones can be localized in a room even when they do not have onset transients? To what extent do available data support models of the precedence effect that involve inhibition in the auditory periphery, as compared with models such as the plausibility hypothesis that involve more central functions? At this time there are partial answers to these questions, and there are methods by which more complete answers can be obtained.

Simulation of Room Effects for Binaural Listening

CHRISTOPH POESSELT AND JENS BLAUERT

Binaural technology offers quite a few opportunities for applications in the fields of audio engineering and acoustical consulting. As an example, we elaborate on the use of binaural methods as tools for room acoustics planning. A three-stage modeling process is proposed:

1. *Sound-specification phase:* In this phase, dummy-head recordings from real rooms and from prior computer simulations are used to help us to determine from the client the sound to be created in the new facility. The recordings to be used in this phase are edited electronically by means of binaural processing algorithms.
2. *Design phase:* During this phase, relatively simple computer models of the planned sound field are created. These models can still be modified with a relatively small amount of effort. The models allow listening to be done binaurally, thus enabling clients and consultants to check against the design goals established during phase one.
3. *Work-plan phase:* This phase goes along with the final specification of the work plan. Detailed computer models or physical scaled-down models of the planned space may be used to decide on details and final adjustments. Tools have been developed to enable binaural listening using such models.

We shall describe the necessary equipment as developed in Bochum for implementation of the planning methods mentioned above. Special emphasis is put on the description of the methods that are used to enable binaural listening using computer models of rooms that are still on the drawing board.

Sensorimotor Adaptation

RICHARD HELD AND NATHANIEL I. DURLACH

Ideally, the design of virtual environment or teleoperator systems must take account of the possibilities and constraints associated with sensorimotor adaptation. In certain cases, performance may be limited by an inability to adapt to unintended transformations of the sensorimotor loop that arise from practical constraints on system construction. In other cases, considerable effort and expense may be wasted developing complex systems to match natural transfer functions when a simple system would provide equivalent performance after only modest amounts of adaptation. Systems that incorporate the detailed transfer function of the human pinna might well constitute an example in this category. In still other cases, it may be possible to include transformations in the systems that are unnatural but that lead to superior performance after adaptation. Thus, continuing with the above example, it might be beneficial to develop an auditory virtual environment in which sensitivity to vertical angle is increased (and made possible at low as well as high frequencies) by incorporating a pinna transfer function that is modeled after a pinna that is much larger than the normal human pinna. Clearly, the benefits of such an unnatural super pinna system would depend crucially on the extent to which the operator can adapt to the new transfer function.

In this paper, we present a brief overview of past work on sensorimotor adaptation, comment on some of the major issues not yet resolved, and consider a number of new research projects relevant to the design of virtual environment and teleoperator systems involving auditory localization.

Multisensory Model for Active Adaptation

GREG L. ZACHARIAS

A model is presented to help account for multisensory active adaptation in structured tasks. A baseline model is first described to account for skilled human performance in dynamic, multisensory, closed-loop tasks; aircraft flight control is used to illustrate a specific model application. The model is then partitioned into parametric components associated with the human's sensory/motor capabilities/limitations, and into structural components associated with the human's internal model of the external task/system characteristics (Figure 1). Parametric adaptation is illustrated via model matches to a single modality (visual) precision tracking task; experimental results show how measured skill acquisition over time is accounted for in terms of inferred parametric adaptation trends (Figure 2). Structural adaptation is implemented via a supervisory loop added to the baseline model; multisensory cueing residuals are introducted to formalize the expected versus the experienced difference and to drive the basic adaptation logic that forces structural changes in the human's internal model of his or her environment. Potential model applications are outlined, including the design of localization/lateralization experiments, using an active psychophysics paradigm, as well as the development of multisensory virtual environments for structured task simulations.

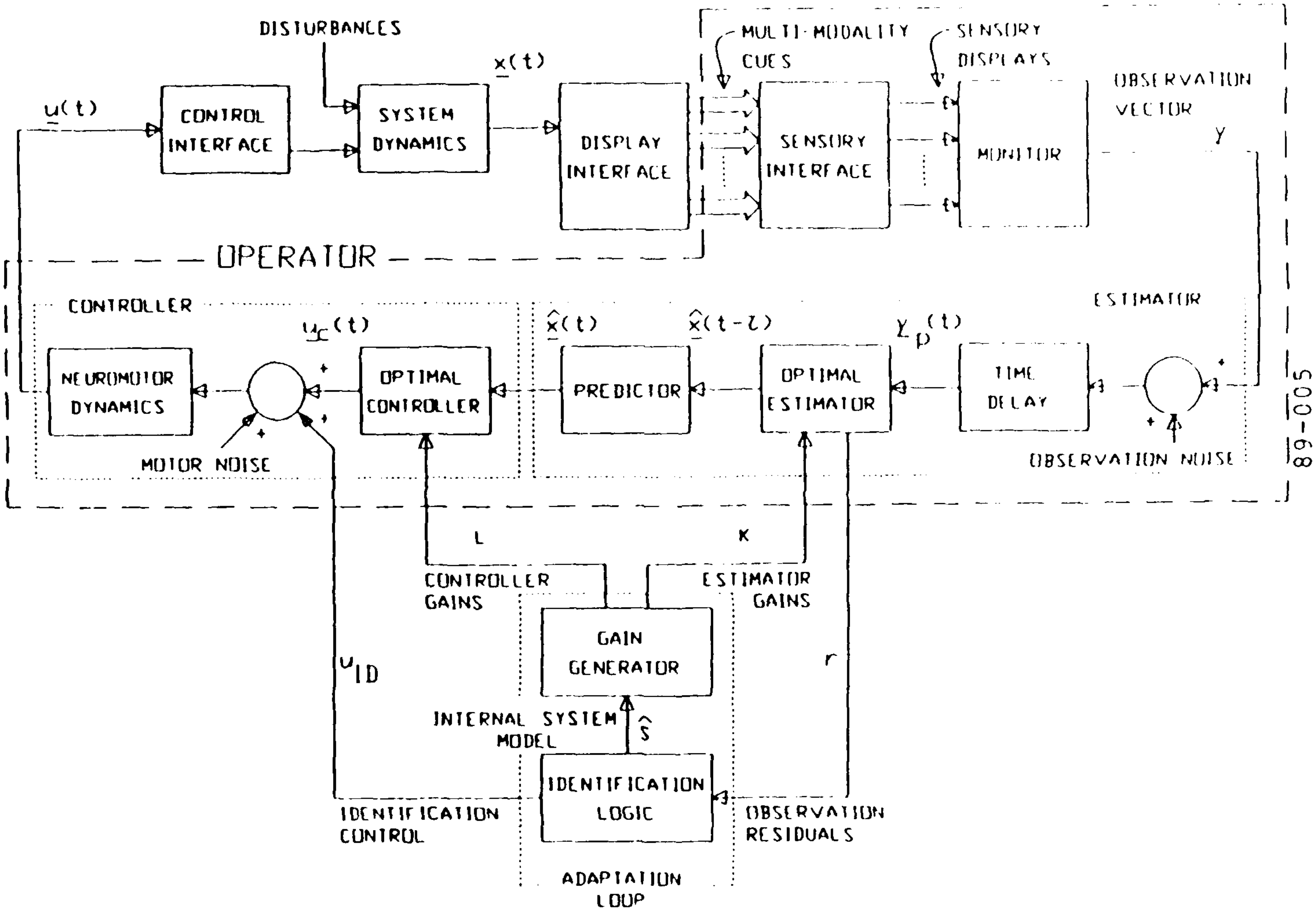
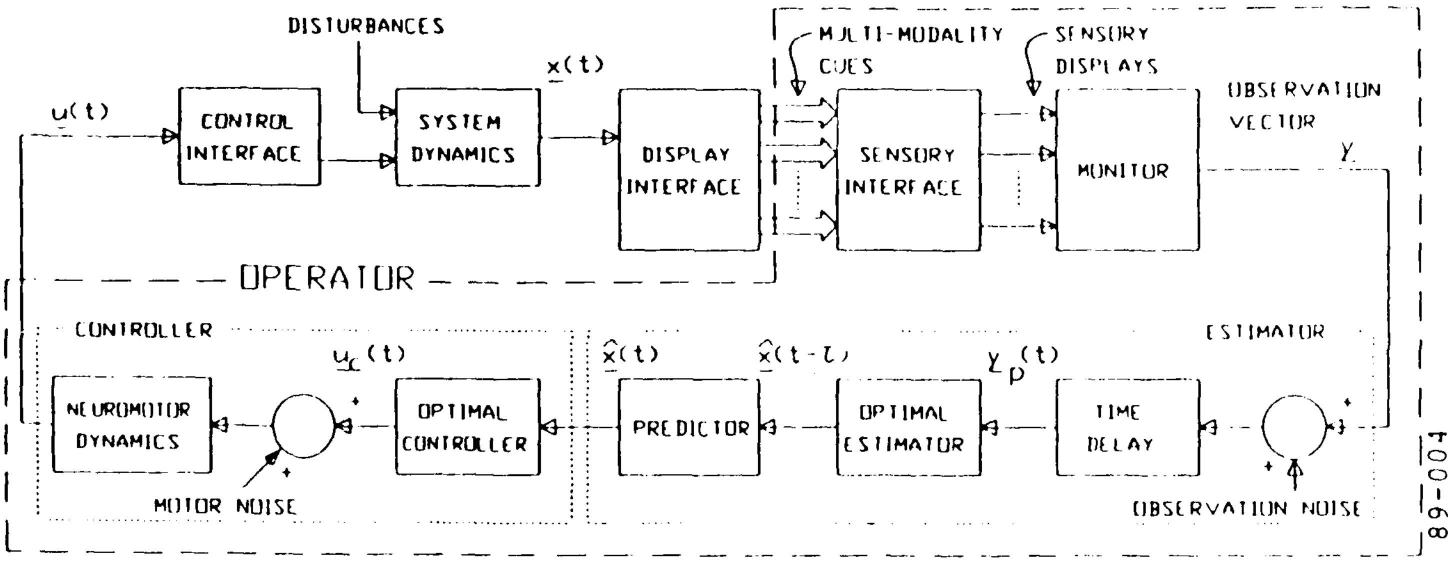


FIGURE 1 Structural adaptation: formation of an internal model.



- REDUCE OBSERVATION NOISE & MOTOR NOISE
- REDUCE TIME DELAYS
- INCREASE MOTOR BANDWIDTH

FIGURE 2 Parametric adaptation: reduction of skill deficits.

Control of Localization in Music Recording and Reproduction

ROBERT BERKOVITZ

Professional concern in the audio industry has shifted from low noise and spectral accuracy, both now easily achieved, toward the task of recording or synthesizing the spatial characteristics of a sound field. New microphones, new microphone configurations, and elaborate signal-processing methods are being used to create the auditory images presented by stereophonic recordings. Proposals have been offered for systems that might reproduce the entire acoustic field in which the listener is placed, or that might at least suggest that such reproduction is taking place. A parallel trend during the past few decades has been the design of loudspeakers and signal-processing devices for home use, some employing techniques well-known in the acoustics literature and claimed to enhance or stabilize the apparent locations of sound sources. The very large audience of listeners now using headphones with pocket cassette tape players has created the possibility of widespread application of binaural recording for music reproduction.

These trends correspond to a shift in economic importance in the music industry from classical, auditorium-based music, with its restricted spatial context, to ephemeral music with a strong emphasis on novelty and sensation. The spatial effects of modern recordings are typically created by assembling a series of individual recordings of discrete sources in a synthetic auditory space. There has been considerable effort to develop analogous methods for systematic control of localization in cinemas, and this has had some effect on music recording. Digital signal-processing systems to provide spatial control in the studio are becoming more complex and powerful, and elaborate digital signal processors for home use are also appearing. These developments are transforming music by altering the expectations of listeners.

Enhancement of Human Localization and Source Separation

PATRICK M. ZUREK

The acoustic signals applied to a person's ears can be thought of generally as a two-channel information display. In normal use, the source of information for the display is the sound field, which is sampled directly at two points. In the case of auditory prostheses, the information source is also the sound field at one or more points, but the sampled signals are transformed by the device to improve the reception of acoustic information by the impaired listener. In the case of some aids for the blind, the information source is also the sound field at one or more points, but the sampled signals are transformed by the device to improve the reception of acoustic information by the impaired listener. In the case of some aids for the blind, the information source is the optical field, and the sampled signals are transformed into an auditory display signal to supplement (or substitute) impaired vision. Future development in areas such as teleoperator systems and human-computer interfaces, as well as prosthetic devices, will increasingly involve auditory display of information.

There are many considerations in the design of auditory displays, and these will vary with the application. One feature of primary importance in normal binaural hearing is the ability to locate images along perceptual dimensions associated with physical space, thus allowing some degree of source separation. Designers of auditory displays would do well to take advantage of this natural means of reducing interference among signals.

This review discusses work of two types. The first type discloses the ability of normal listeners to separate multiple sources, as measured usually with tests of speech intelligibility. This work includes the familiar noise reduction capability of binaural hearing (which will be compared with the performance achievable with fixed and adaptive microphone arrays) as well as the ability to monitor multiple sources while focusing on one. The second type of work concerns systems intended either to exploit or to enhance the binaural ability to separate and localize sources. This work includes the design of supplementary auditory displays that interfere minimally with speech, and the possibility of magnifying interaural cues for enhanced localization and source separation.

BIBLIOGRAPHY

Corbett, C.R.

1986 Filtering Competing Messages to Enhance Mutual Intelligibility. M.S.E.E. thesis. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Durlach, N.I., and X.D. Page

1986 Interaural magnification. *Journal of the Acoustical Society of America* 80:1849-1850.

Kollimeir, B., and J. Peissig
1988 Speech intelligibility enhancement by interaural magnification. *Acta-Otolaryngologica.* In press.

Peterson, P.M.
1987 Using linearly-constrained beamforming to reduce interference in hearing aids from competing talkers in reverberant room. Proceedings of the IEEE International Conference on Acoustics. Speech Signal Proceedings. Dallas, Texas.

Technical Reports Distribution List

Dr. Curt Boylls
National Security Agency
Attn: R523
9800 Savage Road
Ft. George C. Meade. MD 20755

Dr. Allen Boysen
Audiology & Speech Pathology Service
Dept. of Veterans Affairs (126CO)
810 Vermont Avenue, N. W.
Washington, DC 20420

Dr. James E. Densmore, Director
Office of Environment and Energy
AEE-1
Federal Aviation Administration
800 Independence Avenue, S. W.
Washington, DC 20591

Capt. Thomas N. Jones, MSC, USN
Naval Medical Research and Development Command
National Medical Center
Bethesda, MD 20814

LtC. Bruce Leibrecht, USA
U. S. Army Aeromedical Research Laboratory
Attn: SGRD-UAS/LtC. Leibrecht
P. O. Box 577
Ft. Rucker, AL 35330

Dr. Richard Loutitt, Director
Division of Behavioral & Neural Science Room 320
National Science Foundation
1800 G Street, N. W.
Washington, DC 20550

Dr. Andrew Monjan
Neuroscience and Neuropsychology of Aging Program
National Institute of Aging
Building 31, Room 5C-35
Bethesda, MD 20205

Ralph F. Nauton, MD, FACS
Acting Director, Extramural Programs
NIDCD
Federal Building, Room 1C-11
Bethesda, MD 20892

Dr. Arnauld Nicogossian
Office of Occupational Medicine
NASA Headquarters, Code NIH-34
Washington, DC 20546

Dr. Charles W. Nixon
USAF AAMRL - BBA
Wright-Patterson AFB, OH 45433

Dr. John J. O'Hare
Office of Naval Research
Code 1142PS Rm 823
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Christopher Platt
Sensory Systems, BNS, Room 320
National Science Foundation
1800 G Street, N. W.
Washington, DC 20550

Dr. Clemans A. Powell
NASA Langley Research Center
Mail Stop 463
Hampton, VA 23665

Commander
U.S. Army Medical Research and Development Command
Attn: SGRD-PCL/Col. Sedge
Ft. Detrick
Frederick, MD 21701

Dr. Kevin Shepherd
NASA Langley Research Center
Mail Stop 463
Hampton, VA 23665

Dr. Janis Stoklosa
NASA Headquarters/EBM
Washington, DC 20546

Dr. John F. Tangney
AFOSR/NL
Building 410
Bolling AFB, DC 20332

Mr. Richard M. Tedrick, Chief
International and Scientific Program Office, AEE-3
Federal Aviation Administration
800 Independence Avenue, S. W.
Washington, DC 20591

Dr. Thomas Tredici, MC, USAF
USAF School of Aerospace Medicine
Brooks AFB, TX 78235

Dr. Milton A. Whitcomb
NAS-National Research Council
CHABA
2101 Constitution Avenue, N.W.
Washington, DC 20418

Dr. Carl E. Williams
Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, FL 32512

Col. John Wolcott, USAF
Deputy Commander for Research, Development, & Acquisition
Aerospace Medical Division
Brooks AFB, TX 78235

Dr. James W. Wolfe
NASA Suite 600
600 Maryland Avenue, S. W.
Washington, DC 20024

Capt. James N. Woody, MC, USN
Commanding Officer
Naval Medical Research and Development Command
NMCNCR
Bethesda, MD 20814-5044